



universität
wien

DIPLOMARBEIT / DIPLOMA THESIS

Titel der Diplomarbeit / Title of the Diploma Thesis

„I can't even' – how tags are used as communicative
tools on Tumblr“

verfasst von / submitted by

Elisabeth Altmann

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of
Magistra der Philosophie (Mag.phil.)

Wien, 2017 / Vienna, 2017

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on
the student record sheet:

A 190 344 299

Studienrichtung lt. Studienblatt /
degree programme as it appears on
the student record sheet:

Lehramtsstudium UF Englisch und UF Psychologie und
Philosophie

Betreut von / Supervisor:

Univ.-Prof. Mag. Dr. Barbara Seidlhofer

Acknowledgments

Prefacing this thesis I would like to express my deep gratitude to my supervisor Professor Barbara Seidlhofer, who has not only encouraged me to explore the world of ELF, but also been a superb supervisor. I could not be more grateful for her time, helpful inputs and advice, and her willingness to take me on as a supervisee in the first place. Thank you for your enthusiastic and supportive supervision!

My sincerest thanks also go to the Tumblr users who have agreed to participate in this project and shared their insightful observations with me. Thank you for your time and thoughts!

Table of contents

List of abbreviations

1. INTRODUCTION	1
2. THEORETICAL BACKGROUND AND RELEVANT CONCEPTS	2
2.1. AN OUTLINE OF COMPUTER-MEDIATED COMMUNICATION (CMC) RESEARCH	2
2.1.1. FUNDAMENTAL DEFINITIONS	2
2.1.2. CMC beginnings and research interests	4
2.1.2.1. What kind of language is used in CMC?	5
2.1.2.2. Which linguistic features are typical of CMC?	7
2.1.2.3. Does CMC constitute a threat for 'standard language'?	10
2.1.2.4. The prevalence of English in CMC research.....	11
2.2. AN OUTLINE OF ELF RESEARCH.....	12
2.2.1. Fundamental definitions.....	12
2.2.2. ELF research interests.....	16
2.2.3. Key concepts and practices in ELF communication.....	17
2.2.4. English, the Internet, and ELF	24
2.3. AN INTRODUCTION TO TUMBLR	27
2.3.1. What is Tumblr?.....	27
2.3.2. General layout and functioning	29
2.3.3. Tags on Tumblr.....	32
2.3.4. Personal Tumblr experience	33
2.4. TUMBLR IN RELATION TO CMC AND ELF RESEARCH INTERESTS AND CONCEPTS	33
3. RESEARCH DESIGN	41
3.1. RESEARCH METHODOLOGY	41
3.2. ETHICAL CONSIDERATIONS	43
4. ANALYSIS AND FINDINGS.....	45
4.1. RQ 1: THE COMMUNICATIVE NATURE OF TAGS	45
4.1.1. The various kinds of tag usage on Tumblr	45
4.1.1.1. Organization and structuring of content	46
4.1.1.2. Expressing emotion	48
4.1.1.3. Adding commentary.....	51

4.1.1.4. Relating a post to oneself	57
4.1.2. Bloggers' reasons for such tag usage.....	59
4.1.3. Formulaic expressions as notable feature of tag usage.....	62
4.2. RQ2: ELF COMMUNICATION IN TAG USAGE ON TUMBLR	64
4.2.1. ELF on Tumblr	64
4.2.2. Bloggers' reasons for using ELF	64
4.2.3. ELF key concepts reflected in language use in tags.....	67
5. CONCLUSION AND RESEARCH PERSPECTIVES	72
REFERENCES	77
APPENDIX	85
LIST OF FIGURES	85
ABSTRACT	86
ABSTRACT (GERMAN).....	87

List of abbreviations

AoIR	Association of Internet Researchers
CMC	Computer-Mediated Communication
CMD	Computer-Mediated Discourse
CMDA	Computer-Mediated Discourse Analysis
CofP	Communities of Practice
DCOE	Discourse-Centered Online Ethnography
EFL	English as a Foreign Language
ELF	English as a Lingua Franca
ENL	English as a Native Language
L1	first or native language
NS	native speaker
NNS	non-native speaker
WWW	World Wide Web

1. Introduction

When it comes to communication and language use, the Internet is an impressive and extensive breeding ground for diverse and hybrid language behavior. It is therefore not surprising that linguists from a multitude of different fields are interested in the various language phenomena to be found in and across different Internet contexts. The present thesis aims at contributing to this growing field of Internet research by investigating a linguistic phenomenon occurring on the blogging website Tumblr. The linguistic phenomenon in question is the use of tags as communicative tools, in particular as practiced by English as a lingua franca (ELF) users. How this research interest relates to the various chapters of this paper is addressed in the following.

Every time a new medium has been developed, people have been quick to use it for communicative purposes. This is no different when it comes to the Internet, which connects people from all over the world (if they happen to have an Internet connection). Although a lot of different kinds of data and information can be shared through the Internet, one of its main uses in our daily lives is for communicative purposes. Computer-mediated communication (henceforth CMC) provides the context for this thesis and is thus addressed first. Since the linguistic phenomenon of interest here takes place in a CMC setting, namely on the World Wide Web and more precisely on the blogging website Tumblr, some theoretical background knowledge regarding CMC research and its development is considered relevant and thus provided in chapter 2.1. below. Narrowing the research focus down further, the type of CMC investigated is realized through English as a lingua franca (ELF), since it facilitates communication between Tumblr users who speak different first languages and choose to communicate with each other using ELF. Thus, the second part of theoretical background, provided in chapter 2.2., gives an outline of relevant definitions, concepts and considerations in ELF research in order to provide a basis for the analysis and discussion of findings in chapter 4. Specifying the present research even more, the phenomenon is a CMC phenomenon realized using ELF on the blogging website Tumblr. It is thus also of importance to have a general idea of the website and its basic function, which is why Tumblr is presented in chapter 2.3.. The subsequent section then relates some relevant aspects and research interests found in CMC and ELF research to Tumblr and thereby aims at painting a more

comprehensive picture of the website. Within the context of Tumblr, the present thesis addresses the communicative usage of tags on Tumblr and the language chosen to realize it. The following two research questions provide guidance during the analysis:

RQ1: In what ways are tags on Tumblr used to serve communicative purposes?

RQ2: Which ELF key concepts come into play in this type of communication on Tumblr?

The exact research design and ethical considerations are presented in chapters 3.1. and 3.2. With all this background information in mind, the actual analysis and discussion of findings is then provided in chapter 4 and conducted by reference to the two research questions. A final conclusion, containing a summary of findings and suggestions for future research, is provided in chapter 5.

2. Theoretical background and relevant concepts

2.1. An outline of Computer-mediated communication (CMC) research

With the invention and rise of the Internet as a new medium for communication, a new area for research presented itself to linguists. This chapter describes various research interests in the field of computer-mediated communication over the years, which serves as a theoretical basis for the research project at hand. In order to establish clarity about key terms and definitions, these are addressed first.

2.1.1. Fundamental definitions

First and foremost, the term *computer-mediated communication* (henceforth CMC) is addressed. According to Murray (2000: 398), the term was first used to refer to computer conferencing, but has been broadened by researchers to include other communication channels that function via the Internet as well. Murray (P. 1997) points out that CMC “means different things to different people”, which is why it is necessary to be aware of the definition of CMC adopted in any research project. For the present thesis, CMC is understood according to Baron’s broad definition of CMC as “any natural language messaging that is transmitted and/or received via a computer connection” (2003: 10). Not only does this definition include transmissions like e-mails, chat, or messaging on websites, but also transmissions via mobile phone or smart phone connections, since these are computers as well. The first

messages that fit the label CMC were electronic mail messages exchanged between two computers in a shared network in 1972 (Herring 2002: 113). This early mode of CMC, labeled 'e-mail', is still prevalent on the Internet today. In Baron's definition, CMC refers to "a written natural language message sent via Internet" (2003: 10), however, I would argue that since then the notion of CMC has expanded and nowadays is not limited to text formats, but can also be understood to include video and audio messaging that is transmitted via the Internet. The concrete research interest in the present thesis surrounds the phenomenon of tags as communicative tools on the website Tumblr, and is therefore to be situated within the broad context of CMC as delineated here.

As can be seen in the definition of CMC above, the Internet plays an integral role when it comes to this mode of communication since without it, CMC would be impossible. It is thus important to be aware of what the Internet actually is and what it is not. Crystal (2006: 2-3) defines the Internet as:

an association of computer networks with common standards which enable messages to be sent from any central computer (or *host*) on one network to any host on any other.

The Internet originated in the USA in the 1960s and was originally created for the U.S. Department of Defense (Murray 2000: 407), but has since then grown to fulfill a myriad of functions and purposes (Crystal 2006: 3). It is the largest computer network in the world (*ibid.*), connecting millions of people who, according to Herring (2002: 109), mostly use it to communicate via a number of different modes and for a variety of different purposes. Crystal (2006: 24) calls it "an electronic, global, and interactive medium" and argues that these properties affect the language used online, which will be addressed at a later point in this chapter. It is essential to be aware of the fact that the Internet is not synonymous with the World Wide Web (WWW), as Baron (2003: 2) points out, since the Internet predates the WWW and subsumes many functions that are independent of it. The WWW does not entail the entirety of the Internet, but it exists within it, and is defined by Crystal (2006: 13) as

the full collection of all the computers linked to the Internet which hold documents that are mutually accessible through the use of a standard protocol

It was created by Tim Berners-Lee in 1990 to facilitate the sharing of information between physicists, but has spread much further beyond this use and now enables interaction for a plethora of purposes (*cf.* Crystal 2006: 13).

CMC research has dealt with both broad Internet-related phenomena and more individual phenomena related to the World Wide Web specifically. In both cases, and in fact in any kind of CMC research, it is important to be aware of the fact that the electronic environment of the Internet is not a sphere completely separate from our offline lives. As Herring (2002: 152) points out, the juxtaposition of the 'virtual', self-contained world of the Internet with 'real life' is an outdated fiction since Internet use today "is increasingly a part of everyday routine for large numbers of people around the world" (ibid.). This intricate link of the Internet and our everyday lives on the one hand makes it all the more interesting and relevant as a subject of study, but on the other hand comes with the need for increased awareness of assumptions and biases when researching the Internet.

2.1.2. CMC beginnings and research interests

Although the first emails were sent in the 1970s, research into CMC only gained momentum and became a relatively active field of study in the 1990s with questions of classification and characterization of CMC features (Herring et al. 2013: 3). Following this, the focus of attention shifted to studies of CMC discourse in specific textual environments or modes (ibid.). In the mid-1990s, researchers became more and more interested in sociolinguistic aspects of CMC and by then increasingly turned away from the idea of classifying CMC as a whole and instead turned to investigating and classifying specific modes or mode-independent facets of CMC (ibid.). Androutsopoulos (2006, 2008) makes a similar distinction and differentiates between two 'waves' of CMC research. While the first wave focused on "features and strategies that are (assumed to be) specific to new media" and how the new technologies affect language (Androutsopoulos 2008: 1), the second wave comprises pragmatic, sociolinguistic and discourse studies of CMC (ibid.) and thus pays attention not only to the form of language but also to the function and context, as well as its diversity.

It is important to note that most of the CMC research up until the 2010s mostly focused on English-language CMC, which will be addressed again at a later point. For now, in order to gain an understanding of CMC research conducted so far, the following sections provide an overview of the main research interests and insights of the past decades.

2.1.2.1. What kind of language is used in CMC?

As mentioned above, early research into CMC focused on classifying the language used in CMC. Baron (2003: 4) argues that, to a large extent, language on the Internet is natural language and thus, many of the issues around its usage are “not unique to the medium” (ibid.). Nevertheless, researchers such as Baron (2003) and Crystal (2006) have tried to characterize language on the Internet. One of the first questions posed was the categorization of CMC in terms of spoken versus written language (Muniandy 2002). Baron (2003: 4-5) notes that while a large amount of CMC is written, users report it feeling more like spoken language. Indeed, early on research already indicated that CMC can not simply be classified as traditional writing just on the grounds that it is often transmitted as written language and acknowledged that it includes characteristics of spoken language as well. As Squires (2010: 462) notes:

Early CMC scholars were intrigued by the fact that CMC usage seems to encompass canonical features of both speaking and writing: it may be informal, synchronous, and ephemeral—“like speech”—and/or editable, text-based, and asynchronous—“like writing.”

Some CMC aspects related to writing are the transmission modality (Baron 2003: 4), a monologic nature whereby the writer does not know the entirety of their audience in advance because there is no telling how far the text will reach (Baron 2003: 6), decontextualization, and the opportunity to edit a text before it is published (ibid.), as well as the (theoretical) opportunity to be anonymous (Baron 2003: 11), and typically written features like abbreviations or acronyms (Baron 2003:21) or textual formatting (Yates & Orlikowski 1993). Crystal (2005: 1) also mentions the absence of simultaneous feedback and tone of voice, and Herring (2004: 338) notes that language on the Internet leaves a textual trace, which means it can be analyzed and reflected upon more than spoken language can. Murray (2000: 402) also mentions the absence of visual paralinguistic and nonverbal cues. Some CMC aspects related to speech are a dialogic nature wherein multiple participants can be involved (Baron 2003: 6), casual registers that can also be found in, but are not limited to, spoken language (Murray 2000: 397), informality (Yates & Orlikowski 1993), or the expectation of receiving an (almost) immediate response (Crystal: 2006: 28). Crystal additionally lists spontaneity, a loose structure, and social interactivity (Crystal 2006: 42). It thus becomes clear that a simple categorization of CMC as either a form of speech or a form of writing is not achievable, but that CMC seems to incorporate aspects of both modes of communication.

A related distinction is the distinction between synchronous (communicating in real-time) and asynchronous (not communicating in real-time, i.e. with a delay) communication (cf. Baron 2003: 12). While written language typically entails asynchronous communication, speech is a synchronous form of communication. However, CMC does not fit neatly into this dichotomy since different modes of CMC lie at different points on a spectrum ranging from asynchronous to synchronous communication. An email, for example, is asynchronous whereas chat is typically much closer to synchronous communication, although there is likely still at least a minimal delay, depending on the bandwidth of the users' Internet connections.

There are even more features of CMC that do not fit neatly into the written-spoken language divide, such as certain stylistic and linguistic features for example (which will be addressed in the next chapter). Thus, many researchers emphasized the in-betweenness or hybrid nature of language in CMC (cf. Squires 2010: 461). Such is the case in Muniandy (2002) or Crystal's (2006) work. Muniandy looks at CMC, or, in her terms, 'e-discourse', in relation to spoken and written discourse and first makes it clear that while spoken and written discourse differ in terms of contexts, the functions can be the same (2002: 47). Moreover, Muniandy emphasizes that spoken and written texts are not the opposites of a dichotomy, but rather the two ends of a continuum, where features may overlap (ibid.). E-discourse, according to Muniandy, comprises features of both the spoken and written mode and in addition to that has distinctive features which result from it being a new discourse type since it does not rely on typically spoken or written media, but is transmitted electronically (2002: 49). Similarly, Crystal (2006: 1) claims that "Many of the expectations and practices which we associate with spoken and written language [...] no longer obtain" and introduces the concept of 'Netspeak' as a variety of language in CMC that "tries to be like speech" but "remains some distance from it (2006: 41). Crystal, too, argues for the necessity of seeing the distinction between spoken and written language not as a dichotomy but as a continuum and states that different Internet contexts can be found at different points of said continuum (2006: 47). Netspeak, then, displays features of both spoken and written language but according to Crystal (2006: 47-48) also comprises new features that cannot be found along this continuum, which qualifies it as a new variety of language. This is heavily criticized by Dürscheid (2004: 145) who points out two flaws in Crystal's proposition: Firstly, many of the features that Crystal argues to be unique features of Netspeak also occur in different contexts,

and secondly, if there are features unique to the Internet, they are not typical and characteristic features of Internet language in general, but in fact depend on the situational context and its participants. There is thus, as Dürscheid (2004: 154) argues, no general, uniform language use on the Internet, which is why she proposes to abandon macro level research about language on the Internet in general in favor of micro level research which focuses on specific instances of language use instead. On a related note, Squires (2010: 463) argues against the notion of linguistic determinism (2010: 461), which states that linguistic forms are determined by the medium, and points out that not just the medium, but also the format and genre as well as the social characteristics and relations between its users influence online language use.

What this means, then, is that ‘What kind of language is used in CMC?’ is the wrong question to ask, since there is no unified language use across all uses and channels of CMC. Nevertheless it was a question frequently asked in early CMC research, potentially due to the general public’s concern around its possible influence of ‘standard language’. This is addressed in more detail in chapter 2.1.2.3. Related to the concerns about CMC’s influence on any kinds of ‘standard language’ was the investigation of typical linguistic features in CMC, which is addressed in the next section.

2.1.2.2. Which linguistic features are typical of CMC?

While the previous section has shown that a general characterization or classification of CMC is not appropriate, the current section nevertheless provides an overview of stylistic and other linguistic features frequently found in and across several modes or channels of CMC. Early on, researchers identified various linguistic features that they deemed typical occurrences in CMC contexts. They often list microlinguistic features such as divergent typography and orthography or the attempt to make up for the lack of visual or situational cues typical for face-to-face conversation as defining CMC characteristics. Among the first features to be mentioned were emoticons (a combination of punctuation marks and/or numerals to represent certain emotions) and the phenomenon of flaming (a hostile interaction, often characterized by use of vulgar or rude language) (Baron 2003: 20). Baron (ibid.) claims that these features emanated from CMC’s conversational nature and sense of immediacy. Another significant and frequently occurring feature in CMC mentioned by Baron (2003: 21) is

the use of acronyms and abbreviations, which are used due to space, time, or energy constraints or to indicate group membership (ibid.). Murray (2000: 402) mentions simplified syntax, accepting surface errors, using symbols such as multiple vowels or consonants to convey emotion (e.g. as in *stooooooooopp*) or using formulaic phrases (such as emotes) as additional strategies used in CMC to both conserve time and make up “for the lack of paralinguistic and nonverbal cues” (ibid.). In addition to emoticons, non-standard spelling (such as asterisks for emphasis, eccentric spelling, or shortenings) and abbreviations, Bieswanger (2013: 464) also mentions non-standard punctuation, which ranges from a total lack of punctuation to exaggerated use (2013: 476). However, Bieswanger is wary of labeling these features as ‘characteristic’ for CMC and, in line with other researchers like Dürscheid (2004) or Squires (2010), emphasizes that

the occurrence of these phenomena varies according to the language and script used, the mode of CMC employed, the context of the interaction, and other use- and user-related factors (Bieswanger 2013: 464)

Crystal (2006), however, does not shy away from claiming that there are characteristic CMC features and lists all of the features mentioned above as “main linguistic features” of Netspeak (2006: 81-93). In addition to these graphological features, Crystal also points out the existence of an exclusive Internet lexicon, which is made up of various neologisms or new compounds (2006: 82-84). While Crystal concedes that a characterization of Internet language as a whole is impossible, he does so on the grounds of “the transience of the technology” (2006: 224), which again reveals Crystal’s deterministic perspective addressed in the previous chapter. Despite the fact that notions of an ‘Internet language’ like Crystal’s Netspeak have been criticized (see previous chapter), there are still some researchers who reiterate such characterizations (see, for example, Sun 2010). The larger part of CMC researchers, however, started to focus more and more on CMC discourse in different contexts (Herring, Stein and Virtanen 2013: 3) and investigated sociolinguistic concerns as early as the 1990s in what Androutsopoulos (2008: 1) terms the ‘second wave’ of CMC research. According to Androutsopoulos (2006: 421), this

growing body of research inspired by sociolinguistics and discourse analysis [...] aims at demythologizing the alleged homogeneity and highlighting the social diversity of language use in CMC

He further notes that this represents a rejection of technological determinism and shift from “medium-related to user-related patterns of language use” (ibid.) and adds:

Characteristic features of ‘the language of CMC’ are now understood as resources that particular (groups of) users might draw on in the construction of discourse styles in particular contexts. (Androutsopoulos 2006: 421)

In other words, the second wave of CMC represents a shift from describing and attempting to categorize a concept of a general Internet or CMC language to exploring language use in CMC in specific contexts and in relation to various discourses. As Androutsopoulos (2006: 421) points out, this significance of computer-mediated discourse (CMD) and rise of computer-mediated discourse analysis (CMDA, Herring 2001, 2004) goes hand in hand with an interest in social parameters and concepts such as community and identity. A prevalent research focus in CMD research are mode-focused approaches, which addressed CMC within the various contexts of different CMC modes such as email, Internet Relay Chat, or online forums (Bieswanger 2013: 466). Maiz-Arevalo and Garcia-Gomez (2013: 736) mention an interest in different online genres and list emails, chats, instant messaging, weblogs, newsgroup discussions, virtual worlds, video-sharing websites, social networks and even banners and spam as genres that have been studied. When it comes to genres, however, Herring et al. (2013: 9) point out that Internet genres are multi-functional and changeable and thus their definitions have to be equally flexible and open to change (ibid. 10). As a result, very broad definitions have yielded to more pragmatic, context-related definitions (ibid.). By abandoning the idea of CMC as homogenous (Bieswanger 2013: 466) and instead focusing on the differences in language use in different modes and contexts, these studies foster a more detailed and complex picture of CMC. The keyword here is language in use, or discourse (as defined by Widdowson 2007: 7) and it brings with it an extension of research interests, perspectives, methods, and overlaps with other fields. Extending the area of interest from pure text (language used to refer to something; Widdowson 2007: 6) to include discourse, a concept that comprises both “what a text producer meant by a text and what a text means to the receiver” (Widdowson 2007: 7), considerably broadens the field. That is not to say that linguistic features are no longer of interest, but ever since the ‘second wave’ of CMC research, to use Androutsopoulos’ (2006) term, linguistic features are not the sole focus anymore and are put in relation to their context and use. Due to space constraints an overview of

studies and research projects cannot be given here, but some analyses and findings relevant for the research project at hand will be discussed in chapter 2.4.

2.1.2.3. Does CMC constitute a threat for 'standard language'?

The Internet is of tremendous importance when it comes to communication nowadays. Indeed, some (such as Posteguillo 2002: 24) argue that its importance is of equal magnitude to the invention of the printing press in the 15th century due to the fact that it enables human connection and the spreading of information to an extent that was unheard of before. While this is an incredible feat, new technologies are often met with certain wariness and insecurity. This becomes apparent when looking at reactions to the new technology in mass media, as Thurlow (2006) does in his study. Thurlow (2006: 667) points out that mass media exaggerate and caricature language practice in CMC and present it as much more deviant from non-mediated language use than it actually is. As a result of anxieties about negative effects on language use brought about by new technologies, public discourse:

is typically marked by attempts to control the course of language change through the proscription of disfavored forms and the prescription of familiar ones (Thurlow 2006: 668).

Or, as Crystal (2011: 3) puts it:

The prophets of doom have been out in force, attributing every contemporary linguistic worry to the new technology, and predicting the disappearance of languages and a decline in spoken and written standards

Mass media spread negative attitudes related to CMC and CMD (computer-mediated discourse) by propagating themes like the linguistic distinctiveness of CMD or the rise and spread of CMD, by creating a moral panic surrounding it, or by fetishizing CMD practices (Thurlow 2006: 672-679), often without any valid ground for their claims. As Squires (2010: 459) points out, the mass media's practices when it comes to CMC contribute to the enregisterment of Internet language, whereby it becomes differentiable from other registers and gets attributed certain social meanings and (often negative) connotations. An important aspect of enregistered varieties is their connection to language ideologies, as Squires explains (2010: 460). When it comes to Internet language, it is usually contrasted with dominant 'standard language' ideologies, which generally value correctness and disfavor language varieties that do not adhere to the same norms (Squires 2010: 460). This deviance of CMC from

'correct' 'standard language' is what mass media tend to exaggerate and frame as threat to the language caused by the new medium. On a related note, Murray (200: 406) emphasizes "No technology is neutral or value free", which consequently means that any use of computers and by extension CMC carry social meanings, values, and practices (cf. *ibid.*). What follows from this is that CMC is not inherently good or bad, but can be framed in different ways depending on the prevalent ideologies and practices of a certain group. Thurlow considers negative feelings adults have towards CMC and CMD as the result of concerns regarding youth, technology and language, which create a "'triple whammy' panic about declining standards of morality and the unwinding of the social fabric" (Thurlow 2006: 671). While he concedes that media is not academia, he argues that it very much influences and shapes popular understanding and public discourse (2006: 689), which means that its unfounded claims about online language use can be all the more harmful. Unfortunately, moral panic and worries related to CMC and CMD are not likely to disappear anytime soon, as Crystal (2011: 6) predicts, even though "we invariably find they are based on myths" as he claims (*ibid.*). To answer the question posed in the title of this section, then, CMC does not automatically represent a threat, but can, and certainly has been, framed that way by mass media.

2.1.2.4. The prevalence of English in CMC research

By definition, CMC research is concerned with communication and consequently with language. Early CMC research did not, however, concern itself with various different languages in CMC but instead showed a strong English-language bias (Bieswanger 2013: 468), the reasons for which are addressed in more detail in chapter 2.2.3. Research that was published in English seemed to mostly focus on varieties of that language, and while CMC research in and about other languages existed, it received less attention in the early days (*ibid.*). Given the origin of the Internet as an US-American invention, it is hardly surprising that English has long been its most prevalent language. However, many researchers failed to even take other languages used in CMC into account or clarify that they were only referring to English CMC instances in their research (Bieswanger 2013: 468). While it makes sense to closely study English CMC because it represented (and still represents) a large amount of CMC data, conclusions drawn from these studies cannot be applied to CMC of any and every language overall, which is a limitation that has to be kept in mind. An

understanding of which language in particular is the focus of a study is essential for said study's validity and generalized claims bear little value, since they do not necessarily apply to linguistic behaviors in general and across different languages. When looking at CMC research, in particular when dealing with early CMC research, it thus has to be kept in mind that it might only be informative and meaningful in relation to English, if not otherwise stated. At this point, it shall therefore once again be made explicit that the present study concerns itself with a language phenomenon that was observed and analyzed within the context of English as a lingua franca language use on Tumblr. Any findings are therefore not automatically applicable to other languages and their use on Tumblr, although the existence of the phenomenon in other languages is not ruled out either.

Although it seemed an implied standard to investigate English language use in early CMC, Bieswanger states that by now other languages also receive researchers' attention and an increasing number of studies addresses non-English language use in CMC (Bieswanger 2013: 468). In many cases, these are not published in English, however (ibid.), and thus not really accessible for researchers who are not competent in the language in question. Therefore, while there is a diverse growing body of CMC research, the various insights are not always made available to the majority of researchers. One useful tool when it comes to communicating and sharing information with speakers of differing languages is a lingua franca, a contact language between those speakers. English as a lingua franca and its relevance for CMC is the focus of the subsequent chapter.

2.2. *An outline of ELF research*

2.2.1. Fundamental definitions

As with almost any discussion of theoretical background it seems most purposive to start a discussion of English as a lingua franca (henceforth ELF) by considering its definition(s) and its most relevant related concepts. Thus, the present section aims at introducing the understanding of ELF assumed for the present research by means of reviewing some significant ELF research interests.

Although lingua francas, or "contact languages used among people who do not share a first language" (Jenkins, Cogo and Dewey 2011: 281) such as ELF have been around for hundreds of years, Jenkins et al. state that ELF as a research

interest is much younger (ibid.). Starting out as language teaching-related research into formal aspects of ELF in the 1980s, research into ELF as it is used, and adapted in specific communicative situations did not become prominent until the new millennium (Jenkins et al. 2011: 282) with applied linguistics taking an interest in it (Jenkins et al. 2011: 282-283). The growing awareness of English as a global language used in a number of different global contexts (cf. Crystal 2003) has likewise led to more and more attention being paid to the phenomenon of English as a lingua franca in various and varied contexts (Jenkins et al. 2011: 284) and necessitated close examination of how ELF is to be understood and defined. For the present thesis, the following definition by Seidlhofer (2011: 7) is considered the most suitable and inclusive:

I therefore prefer to think of ELF as *any use of English among speakers of different first languages from whom English is the communicative medium of choice, and often the only option.* (Seidlhofer 2011: 7, emphasis in original)

What makes this definition more inclusive compared to other ELF definitions is the fact that Seidlhofer does not exclude English native speakers from the group of ELF speakers. She does, however, note that they constitute a minority of ELF speakers and “their English will therefore be less and less likely to constitute the linguistic reference norm” (Seidlhofer 2011: 7). Moreover, as Jenkins et al. (2011: 283) note, “ELF is not the same as English as a Native Language (ENL)” and thus has to be acquired by English native speakers (henceforth ENS) and English non-native speakers alike, notwithstanding the fact that ENS are in a somewhat privileged position which facilitates the acquisition process (ibid.). Not all researchers agree that ENS should be considered ELF users. In Firth’s definition for example, ELF is understood as

a ‘contact language’ between persons who share neither a common native tongue nor a common (national) culture, and for whom English is the chosen *foreign language* of communication (Firth 1996: 240; emphasis in original)

However, this much more narrow definition of ELF is only adopted by a minority of researchers, as Jenkins et al. (2011: 283) claim. Thus I turn back to Seidlhofer’s definition. As Seidlhofer emphasizes, her definition represents a functional and not a formal conceptualization of ELF (ibid.), which is to say that it defines ELF in terms of its use, not its formal features. This focus on function over form mirrors the shift of attention in ELF research from ELF features that differentiate it from English as a

native language (ENL) to considering ELF on its own terms and for its uses. In other words, this is a distinction between what Seidlhofer, Breiteneder and Pitzl (2006: 9) call an 'exonormative' approach, which focuses on ELF as differing from ENL and an 'endonormative' approach, which is interested in the process of ELF communication per se. It is important to note, however, that this differentiation, like so many others, does not represent a strict dichotomy but exists in gradients.

Another fundamental distinction relevant when discussing ELF and ELF research is the distinction between 'standard English' and other varieties. Multiple times in this thesis the notion of English has been evoked already in some context or other but what exactly 'English' is meant to denote has not yet been addressed. Thinking of English, we typically have a conception of some kind of 'standard English' spoken by native speakers in mind, but given the fact that most speakers of English nowadays are not ENL speakers but English as a foreign language (EFL) or ELF speakers (Seidlhofer 2004: 209s), this common conception has to be called into question¹. In line with this, Baker calls for:

a conception of English that accepts a plurality of Englishes and an understanding that English is not seen as the property of one culture or community (Baker 2009: 568)

In a similar vein, Crystal considers English to be an effective global language which has both native speakers and non-native speakers but should not unnecessarily emphasize or value this distinction (Crystal 2003: 6). Instead, the focus lies on its power as a language resulting from the power of its people (Crystal 2003: 9). As Crystal argues, English did not become a global language due to any features of the language itself, but due to the economic, technological, cultural, political and military power of the people speaking it (2003: 7, 9) and the history that resulted out of this power. Therefore, conceptions of a somehow superior 'standard English' are not related to any superiority of this standard, but to the dominance of its speakers. But who are these speakers? In times of globalization it is not easily claimed that one group of speakers has more power over English than the other, especially considering that English native-speakers are outnumbered by the language's non-native speakers. It therefore seems unreasonable to consider 'standard English' the 'main form' of English which happens to have a variety of different subforms and, indeed, Trudgill (2002: 160, 165) recognizes it not as *the* standard but as one of the

¹ which is the reason why I always put 'standard English' in quotation marks in this thesis

many varieties of English, in other words a dialect. Accepting this conception of 'standard English' allows for a view of different Englishes that does not indubitably consider 'standard English' as the ideal for each and every English speaker and instead shines a light on the plurality of different Englishes. It is among these Englishes that we find not only English as an international language or English as a foreign language, but also English as a lingua franca, which is the most relevant type of English language communication for this paper. ELF is, however, not a variety of English as Seidlhofer (2011) and later Cogo (2012: 98) claims. Cogo points out that "a variety is the type of language spoken by a precise speech community" (Cogo 2012: 98), which is not a flexible enough definition for the phenomenon of ELF communication. After all, ELF can be both localized and transnational, and it is much more heterogeneous and variable than what is traditionally understood as a language variety (ibid.).

It thus seems more useful not to try to put ELF into the right pre-existing boxes, but consider it on its own terms in regards to different language aspects. Meierkord (2002) investigated ELF in relation to culture and asked herself the question whether it is better understood as a 'language stripped bare' or 'linguistic masala'. Based on the data in her study she came to the conclusion that it is both, and that participants' linguistic behavior "largely depends on what culture a speaker wants to construct in a particular conversation" (Meierkord 2002: 128-129). Correspondingly, Baker (2015: 3) claims that it is "naïve to assume that ELF is a culturally and identity neutral form of communication", since no form of communication can truly be neutral (2015: 37). Therefore, Baker calls for contextualized perspectives when examining intercultural communication like ELF communication if the goal is to attain any kind of understanding (2015: 27). This again indicates that, due to its variability and fluidity, ELF is best understood in regards to each specific communicative situation and context. Comparing it to certain varieties of English, such as 'standard English' for example, does not seem the right basis for any rewarding characterization unless an exonormative approach is adopted (cf. Widdowson 2015).

An introduction to some of the research into the nature of ELF is presented in the next section and aims to provide a more detailed understanding of this global phenomenon.

2.2.2. ELF research interests

The nature of ELF communication has predominantly been studied in relation to spoken language (Jenkins et al. 2011: 286; Seidlhofer et al. 2006: 10), which sets the present research apart from a majority of ELF research since it is concerned with written language. So far, ELF researchers have paid most attention to the aspects of lexis, lexicogrammar, pronunciation and pragmatics (Jenkins et al, 2011: 286), which relates to the interests in intelligibility, accommodation and negotiation of meaning in ELF communicative situations (Jenkins et al. 2011: 287). In this context, Jenkins researched and categorized certain lingua franca core features when it comes to pronunciation, which, when available in ELF speakers' repertoires enhance intelligibility (ibid.). Exonormative approaches like this are not just found in ELF pronunciation research, but also in research into ELF lexis or lexicogrammar. As Jenkins et al. (2011: 288) note, the analysis of ELF data collected in ELF corpora such as VOICE (The Vienna Oxford International Corpus of English, www.univie.ac.at/voice) or ELFA (English as a Lingua Franca in Academic Settings, Mauranen 2003) has led to the identification of certain patterns in ELF lexis and grammar. Despite their high frequency and occurrence within a wide variety of different ELF contexts, researchers have been tentative about identifying and labeling them as 'characteristic' features of ELF and instead consider them communicatively effective ELF variants (Jenkins et al. 2011: 289). Although lexicogrammatical features in ELF communication are thus understood as flexible, some features observed by Seidlhofer (2004) have been shown to be relatively stable (ibid.). These features, which Seidlhofer identifies as certain regularities in ELF communication which would probably be corrected in English language teaching settings but "appear to be generally unproblematic and no obstacle to communicative success" (Seidlhofer 2004: 220), include:

- ‘Dropping’ the third person present tense –s
 - ‘Confusing’ the relative pronouns *who* and *which*
 - ‘Omitting’ definite and indefinite articles where they are obligatory in ENL, and inserting them where they do not occur in ENL
 - ‘Failing’ to use correct forms in tag questions (e.g., *isn’t it?* or *no?* instead of *shouldn’t they?*)
 - ‘Inserting’ redundant prepositions, as in *We have to study about...*
 - ‘Overusing’ certain verbs of high semantic generality, such as *do*, *have*, *make*, *put*, *take*
 - ‘Replacing’ infinitive-constructions with *that*-clauses, as in *I want that*
 - ‘Overdoing’ explicitness (e.g. *black color* rather than just *black*)
- (Seidlhofer 2004: 220²)

Although proposed as preliminary findings, these features provided a starting point for many a study (Jenkins et al. 2011: 290). What is of interest regarding lexicogrammatical features in ELF, however, is not the existence of certain salient features per se, but their functional use and semantic properties, as Seidlhofer (2009b) points out. Another significant contribution to ELF research when it comes to formal variation is Pitzl’s study of idiomatic expressions (Pitzl 2009). In her study Pitzl illustrates how idiomatic expressions in ELF interactions differ from conventionalized ENL idioms due to the fact that they are not as formally bound (ibid.). According to Pitzl, speakers in ELF situations construct idiomatic language by modifying existing English idioms, creating metaphorical images by translating idioms from other languages, or coining new metaphorical images and thus exhibit much more idiomatic freedom than ENL speakers tend to do (Pitzl 2009: 317).

This and other more recent ELF research focused more on function rather than on form represents the movement from normative and form-focused research to function-focused pragmatic research. Much like the context-specific CMC research mentioned in chapter 2.1.2.2, ELF research related to studying pragmatic aspects in a variety of different communicative situations indicates the linguistic diversity of ELF. Nevertheless, some key concepts and pragmatic strategies have occurred again and again in ELF communication. These are addressed in more detail in the next section.

2.2.3. Key concepts and practices in ELF communication

When it comes to the notion of key concepts or ‘normal’ practices in ELF communication, it is important to be aware that these represent abstract, albeit

² Quotation marks reinstated by Seidlhofer in personal communication

useful, distinctions created through our meaning-making processes, as Widdowson (2012: 5) points out. As a result, categorizations and classifications according to these constructed concepts reflect abstractions rather than a genuine breakdown of actual nature. Nevertheless, we need these ‘convenient fictions’ as Seidlhofer (2011: 70-74) calls them to be able to cope with the sheer amount of data present in the world. That being said, Widdowson emphasizes that while we need these constructed simplifications, we also need to be aware of their limits when it comes to their relative validity and convenience (Widdowson 2012: 7-8). With that in mind, the following discussion of ELF key concepts and practices put forth by different ELF researchers is not meant to suggest that all of these features are considered essential and necessary elements of each and every ELF interaction. They are, however, helpful general concepts for approaching a description of ELF communication, and the concepts discussed in the following are adopted as such for the purpose of the present research.

One key concept when it comes to ELF seems to be its fluid nature. Owing to this, ELF interactions are not easily put into neat boxes. As Ferguson among others argues, ELF interactions are of a “fluid, hybrid, heterogeneous and variable” nature (Ferguson 2012: 177) and:

Strict conformity to the norms of some abstract standard language construct is largely irrelevant, beside the point. (Ferguson 2012: 178)

According to Firth (2009: 162) this variability and diversity is “at the heart of ELF encounters” and might constitute what could be labeled the ‘lingua franca factor’, if there is such a thing. As a result, Firth argues that ELF is not easily characterized in separation from its interactions and speakers and is better considered as emergent in communicative situations (Firth 2009: 163). Nevertheless, while there might not be an inherently typical form of ELF communication, some communicative strategies and dominant concepts can be found across a number of different ELF situations.

In lingua franca situations, participants endeavor to achieve effective communication by doing interactional work (Firth 1996: 256). This mainly includes “basic mechanisms of conversation, notably those of turn-taking, sequential relations, and topic management” as Firth (1996: 256) argues. However, in many ELF interactions further principles are at work, since, as has already been established above, ELF interactions differ from ENL and other types of interaction according to their flexibility and variability. Hülmbauer defines two dimensions of

linguistic flexibility in ELF: “(a) the dimension of non-native speaker status and (b) the dimension of interculturality” (Hülmbauer 2013: 55). Considering these wide ranges of action ELF interactants are thus faced with, it is not surprising that they need to negotiate meaning intersubjectively and accommodate each other in order to achieve meaningful interaction. In each communicative situation, ELF speakers may draw on their plurilingual resources (Hülmbauer 2013), but need to ensure mutual understanding by employing certain communicative strategies which go beyond basic language proficiency, as Seidlhofer (2004: 222) argues. A general communicative competence which involves accommodation skills and awareness of shared knowledge and its limits is also needed (*ibid.*). Given the premise of ample variation and flexibility in ELF interaction, it would not seem irrational to expect a certain amount of misunderstandings, yet this supposition has not been validated by research. As Seidlhofer (2004: 218) states:

Misunderstandings are not frequent in ELF interactions; when they do occur, they tend to be resolved either by topic change or, less often, by overt negotiation using communication strategies such as rephrasing and repetition.

Judging from the research, misunderstandings do not seem to be a major problem in ELF interactions, though it has to be noted that this might just be a reflection of the data analyzed. Similarly, as is suggested by the ‘let-it-pass principle’ (Firth 1996: 243), speaker errors or listener uncertainties regarding an utterance are also not a critical issue in ELF communication as long as a basic consensus and understanding are achieved in the end. Related to this collaborative negotiation of meaning is the cooperative imperative, which calls for cooperation and adjusting of language in order to effectively communicate with others (Seidlhofer 2009a: 195). Seidlhofer (2009a: 197) relates this to Sinclair’s idiom principle and to the least effort principle: The availability of idioms or preconstructed phrases enables interactants to communicate without having to make every single semantic choice themselves, which means less effort is needed for communication. In order for these principles to be effective, however, interactants typically need to be aware of the idiomatic patterns accepted by a certain community (Seidlhofer 2009a: 198). In ELF interactions, the patterns deemed acceptable do not need to reflect native speaker patterns but are in fact, as Seidlhofer (2009a: 211) argues, co-constructed by the ELF speakers themselves in order to serve communicative purposes in each specific

interaction. This consideration of norms and their validity among certain groups of people leads to the various concepts of community discussed next.

Terms like ‘community’, ‘speech community’, or ‘community of practice’ are frequently employed, but also challenged concepts, in ELF research. Owing to the fluid and context-dependent nature of ELF, some kind of definition of the group of people certain observations or claims are valid for is inevitable. However, it is not as simple as just labeling everything a community, since this term carries certain notions which need to be taken into account. As Seidlhofer notes:

traditionally, a community has been understood in a predominantly physical, local sense. [...] with the expression “the international community” as a kind of metaphorical extension of this sense. (Seidlhofer 2007: 313)

But over time, and hand in hand with globalization and higher degrees of international connectedness, the primarily physical and local conception of community has ceased to encompass everything commonly considered a community (cf. *ibid.*). Enabled by technological advancements, communities nowadays may consist of members all over the world (Seidlhofer 2007: 314), and, as Seidlhofer (2007: 325) argues “inevitably our sense of what constitutes a **legitimate community** [...] has to change, too” (emphasis in original). In current research, different conceptions of community exist side by side. Speech community, for example, is defined by Posteguillo (2002: 31) as a label for “a group of people who use the same dialectal variety within one language, or simply the same language”. Posteguillo then distinguishes this from the term discourse community, which is used when the community uses a written code for communication (*ibid.*). Another concept of community is Wenger’s notion of a community of practice. Wenger’s conception of communities of practice locates community around a shared venture or interest, which results in some kind of shared sense of identity (Wenger, McDermott and Snyder 2002: 4-5). Communities of practice are always about something (Wenger et al. 2002: 43), and while they may take many different forms, they all feature three basic elements:

a *domain* of knowledge, which defines a set of issues; a *community* of people who care about this domain; and the shared *practice* that they are developing to be effective in their domain. (Wenger et al. 2002: 27; emphasis in original)

The concept of communities of practice (henceforth CofP), as Seidlhofer (2009a: 210) argues, “is well suited to capturing what goes on in many ELF interactions” and moreover, global communities of practice might even be defined by the fact that their community members do not share a native language (Seidlhofer 2007: 314). According to Angouri (2016: 327), the CofP framework is particularly useful in that it provides a basis for studying members’ practices in relation to both local language use but also to broader socio-cultural contexts. However, despite its serviceability, the notion of CofP might not be adequate in every ELF communicative situation, since, as Vettorel (2014: 28) points out, not every grouping of ELF users meets Wenger’s criteria. That is to say, the notion of community adopted in a research project needs to actually apply to the group of people it refers to. In CMC contexts, additional concepts of community can be found. Posteguillo (2002: 32) argues “that Internet is constructed on a set of overlapping speech and discourse communities or communities of practice” and proposes the term netcommunities. He defines a netcommunity as follows:

A netcommunity, then, is a networked community of practice made up of Internet users who share the use of one-or a few-cybergenres for a certain set of common purposes. (Posteguillo 2002: 32)

Herring (2004) addresses the concept of ‘virtual community’ and lists six sets of criteria gathered from literature on virtual communities. These are:

- 1) active, self-sustaining participation; a core of regular participants
 - 2) shared history, purpose, culture, norms and values
 - 3) solidarity, support, reciprocity
 - 4) criticism, conflict, means of conflict resolution
 - 5) self-awareness of group as an entity distinct from other groups
 - 6) emergence of roles, hierarchy, governance, rituals
- (Herring 2004: 351)

Based on these criteria, Herring breaks down the notion of virtual community into “component behaviors that can be objectively assessed” (Herring 2004: 352) and hypothesizes certain discourse behaviors on the levels of structure, meaning, interaction, social behavior, and participation to indicate virtual community (Herring 2004: 356). However, Herring (2004: 359) argues that when it comes to assessing whether a group of people constitutes a community or not, it would be ideal to recognize and consider the participants’ own perceptions. While ethnographic studies provide useful insights into norms and behaviors within certain online environments “in which participants may experience a strong sense of subjective belonging”

(Herring 2004: 362) Herring warns that “the studies do not prove or disprove the existence of virtual community” (ibid.). All in all, different notions of community might be useful for different types of research and research interests, but one has to bear in mind that they are somewhat artificial constructs (cf. Herring 2004: 361, Widdowson 2012: 8). Kramsch (1998: 10), for example, makes use of the construct of “a discourse community that shares a common social space and history, and common imaginings” to define culture, a concept which plays a part in the next section.

Another prominent concept when it comes to ELF research is the concept of identity. As has already been indicated in the previous section, identity is closely related the notion of community. Wenger, McDermott and Snyder (2002) for example used the concept of a common sense of identity to define communities of practice. But what, exactly, is identity? In Baker’s words, subjective, self-constructed identity can be considered as “the web of social groups that we identify with” (Baker 2015: 108), but, as we will shortly see, it has to be kept in mind that not all identities are self-constructed. Norton (2010: 350) claims that identity is something we negotiate and renegotiate anew in every conversation we have with others. A similar understanding of identity as “a relational and socio-cultural phenomenon that emerges and circulates in local discourse contexts of interaction” is taken up by Bucholtz and Hall (2009: 18) who offer a very broad definition of identity as “the social positioning of self and other” (ibid.). Identity is thus an inherently social construction since it can only be negotiated if there is an other. This also indicates that identity is connected to some form of action, since negotiation does not happen on its own. Baker (2015: 42) underscores the need for activity when he states “Identity is, therefore, not something we have but is something we do” and, additionally, “identity is always constructed in relation to others.” (Baker 2015: 112). According to Pavlenko and Blackledge, this construction, offering and validation of identity through discourse represents a social constructionist view of identity (Pavlenko and Blackledge 2004: 14) and emphasizes the role of language. Similarly, in Risager’s (2006) research into language and culture, linguistic practice is considered “as constitution of identity” (2006: 75), although practical aspects are also said to play a role in language choice. Language is of great importance when it comes to identity in CMC contexts as well, as indicated by Danet and Herring’s claim that “language choice and language use are the primary means of signaling cultural identity in text-based CMC” (2007: 7). However, even though language and identity

can, especially in CMC contexts, be played with (Danet and Herring 2007: 12), identity construction is not always necessarily deliberate, as Bucholtz and Hall (2009: 25) note. They state:

Any given construction of identity may be in part deliberate and intentional, in part habitual and hence often less than fully conscious, in part an outcome of interactional negotiation and contestation, in part an outcome of others' perceptions and representations, and in part an effect of larger ideological processes and material structures that may become relevant to interaction. (Bucholtz and Hall 2009: 25)

Baker (2015: 111) and Pavlenko and Blackledge (2004: 21) also argue in line with this and emphasize that not all identities are negotiable. As Pavlenko and Blackledge (2004: 27) note "linguistic and identity options are limited within particular sociohistoric contexts". Accordingly, Baym (2000: 143) argues that what the acceptable means of establishing identity are is socially defined. Even though various definitions of identity have been provided above, it all seems to come back to the fact that identities are related to broader contexts, especially contexts of community and culture.

When it comes to ELF contexts, Seidlhofer and Widdowson (2007: 259-362) argue that the territorial imperative applies and results in expression of both social identity but also group membership through the use of certain slang or idiomatic patterns as territorial markers. The relevance of the reconciliation between individual cultural identity and intercultural intelligibility is particularly evident in lingua franca communication (Pölzl & Seidlhofer 2006: 151). However, as Meierkord (2002: 126) notes, in some ELF interactions (particularly in business contexts) the communicative goals might take precedence over the negotiation of identity. Nevertheless, in principle, signaling of cultural identity is observable in numerous instances of ELF interaction, as is indicated by Klimpfinger's (2007) research into code-switching for example. In CMC contexts, ELF can be and is used to communicate both global and local identities to a large international audience and is thus favored by many users (Vettorel 2014: 16-17). When it comes to new media such as the Internet, it can be seen that shared interests and values may be more relevant for identification than national or local identities (Leppänen, Pitkänen-Huhta, Piirainen-Marsh, Nikula and Peuronen 2009: 1080). Correspondingly, Pennycook (2010) suggests that new media and the flows of culture, especially popular culture, it brings with it constitute an essential site of identity construction and sharing. As will be shown in the next

section, English, and in many cases ELF, plays a pivotal role in this. This once again indicates that concepts like ELF, identity, culture, and community are all interwoven in a web of language, culture, and society and not easily defined as separate entities. It was thus the aim of this section to provide some glimpses at the manifold connections and different conceptions of key concepts related to this research project.

2.2.4. English, the Internet, and ELF

When it comes to the Internet, English and especially ELF plays a prominent role. As has been mentioned before, the Internet originated in the United States and both its successive development and commercial spread predominantly commenced there (Murray 2000: 407). As Murray (2000: 407) states, this technological development, together with the general rise of English as a global lingua franca has significantly influenced the prominence of the English language in CMC contexts. Baron (2003: 8) notices two language trends on the Internet: On the one hand, there seems to be a trend towards the adoption of ELF as *the* Internet lingua franca, and on the other hand, Internet users “produce increasing amounts of content in languages other than English” (ibid.). Thus there is seemingly both a dominance of English but also linguistic diversity on the Internet. Linguistic diversity here does not solely refer to different languages, but also means different variations of the English language itself, since it cannot be claimed that there is only one English on the Internet. Therefore it stands to reason that the notion of ‘English online’ subsumes native speaker and non-native speaker Englishes (such as ELF) alike.

Block (2004) addresses the fear prevalent among some scholars that English might completely dominate other languages when it comes to international endeavors and argues that, even though the Internet was originally geared towards the English language and additional advantages rose out of the restriction of early Internet access, “English never quite became 100% dominant” (Block 2004: 32). As Block (2004: 35) argues, such predictions of dominance are “overly pessimistic” and language diversification is slowly taking root in CMC. Indeed, a look at some relatively recent statistics provided by www.internetworldstats.com shows that the English language has not overpowered all others. According to an Internet World Stats table dated June 30th 2016, only 26.3% of Internet users use English online, which is closely followed by 20.8% of Internet users who use Chinese

(<http://www.internetworldstats.com/stats7.htm>). While Spanish, Arabic, Portuguese, Japanese, Malay, Russian, French and German are listed as other prevalent languages used on the Internet, 22.1% of Internet users, and thus not a small margin, use languages other than those top ten languages (ibid.). Although Internet World States concedes that such data is not easily gathered and represents estimates, they claim that their sources are reliable (ibid.). However, as Paolillo points out, there is no Internet census and Internet user data is often derived from marketing companies, which might implicate a certain bias and thus affects the data's validity (Paolillo 2007: 408). Whether the figures are accurate to a T or not, they are certainly indicative of a diversification of languages used online. As Sebba (2013: 100) argues:

the internet has produced a large additional space, relatively free from normative constraints, in which speakers can practise multilingualism in written, computer-mediated communication

In line with this, Vettorel (2014: 6) claims:

The wider networking spaces allowed by new technologies provide thus significant room for internationally oriented, translocal practices

Despite this freedom of language choice and diversity of language practice as hinted at in the statistics discussed above, Danet and Herring (2007: 22) suggest that English has been reinforced as a favorable language of wider and global communication both online and offline. Paolillo (2007: 425), too, argues that “the current situation favors large languages, and especially English”. As a result, English and ELF in particular are of special interest when considering language choice and language use on the Internet. As Warschauer, El Said and Zohry (2007: 314 in line with Crystal 2003) point out, economic and social globalization have necessitated an international lingua franca and thereby further fostered the global ascendancy of English. However, as is indicated by the studies reported in Wright (2004) English does not always automatically take precedence over Internet users' national languages and, as Androutsopoulos sums up those study results, “reported language choices vary according to communication mode and web content” (Androutsopoulos 2006: 429). It is thereby argued that multilingual Internet users do not simply adopt one language as their form of communication in CMC contexts, but instead utilize their multilingual resources and choose the language according to contextual and content factors. Professional communication seems to be one of the contexts in

which ELF is the preferred choice (Androutsopoulos 2006: 429; Seidlhofer, Breiteneder and Pitzl 2006: 5), but Seidlhofer et al. (ibid.) also mention the importance of English in public domains and emphasize the utility of ELF when it comes to connections based on common interests. They state that “It is in these domains that English has evidently been spreading beyond the elites” (ibid.). Nevertheless, this does not rule out the use of “English as a tool of power”, and Internet users not well-versed in the language may still be excluded from various Internet contexts, as Murray notes (Murray 2000: 408). Such was the case for example when languages not based on the Roman alphabet could not be used and depicted in early CMC due to a lack of related software (Murray 2000: 409). Moreover, as is indicated by Androutsopoulos’ (2013: 68) investigation of code-switching in CMC contexts, website hosts have an influence on their websites’ linguistic environment by providing the communicative setting for CMC interactions. If a website’s overall layout and navigation links and the like are only presented in one language, its users need some basic level of command of that language in order to be able to use the website effectively. Since many CMC contexts such as websites, applications, or software are English-language based, having the necessary English language competence to use them is vital for many Internet users. Elsewhere, Androutsopoulos (2010: 211) proposes four dimensions of language in social networking sites and content sharing sites, and while users may freely choose the language for three of them, the fourth, organizational, dimension is determined by the site designer. Thus, potential users can be excluded and limited in their agency based on their language competences. The language competence needed to effectively use a website might also relate to Internet users’ general preference of websites in their native language, if available (Vettorel 2014: 13).

The various considerations of language(s) and the Internet illustrated above highlight the two conflicting language trends on the Internet stated at the very beginning of this section. When it comes to the CMC context of blogs in particular, both these trends can also be observed, as Schmidt (2007: 1410) indicates when he claims that many studies of blogs focus on English-language blogs but a growing body of research addresses other languages as well. Despite this growing interest in other languages, the focus in the present thesis lies on ELF and bloggers’ reasons for choosing it. Vettorel (2014: 13) suggests that bloggers seem to choose ELF for “internationally-oriented purposes” and in wider networking, that is to say, if they want

to address an international audience. Therefore, English seems to be the dominant language in the blogosphere (Vettorel 2014: 14-15). However, bloggers do not only choose English to achieve a wider reach, but also to indicate their belonging to a particular (English-speaking) online community or culture, as Šabec (n.d.: 9) suggests. According to Šabec, some bloggers thus give preference to their global identity over their local one (Šabec n.d.: 11) and use ELF because it indicates a kind of online prestige that their native language does not retain (Šabec n.d.: 11). Being able to communicate with a diverse international audience, however, is more frequently reported as the main reason for blogging in English (Vettorel 2014: 116). Whether identity expression or a wider global reach are the deciding factors, bloggers deliberately choose to communicate via English (Vettorel 2014: 91, 116). In the analysis in chapter 4, Tumblr users' reasons for choosing ELF will be addressed and related to the reasoning illustrated here.

2.3. *An introduction to Tumblr*

2.3.1. What is Tumblr?

In order to gain an understanding of the phenomenon that will be analyzed and discussed in chapter 4, it is important to be familiar with its context. Therefore, I now describe the blogging website Tumblr, which, in Page, Barton, Unger and Zappavigna's terms, makes up the generic context (Page et al. 2014: 33) and provides the cotext (ibid.) for each use of tags.

Tumblr is a blogging website which allows users to post various kinds of content. On its Login page, Tumblr is described as follows:

Tumblr is so easy to use that it's hard to explain. We made it really, really simple for people to make a blog and put whatever they want on it. Stories, photos, GIFs, TV shows, links, quips, dumb jokes, smart jokes, Spotify tracks, mp3s, videos, fashion, art, deep stuff. Tumblr is 340 million different blogs, filled with literally whatever.

(https://www.tumblr.com/login?redirect_to=%2Fdashboard, last accessed March 16th 2017)

Although it does not label itself as such, Tumblr is sometimes also referred to as a microblogging website (e.g. Bourlai & Herring 2014; Golbeck 2015). Chang et al. describe microblogging platforms as "in between of traditional blogging and online social networking services" (Chang, Tang, Inagaki and Yan 2014: 21) and dictionaries of Internet-related terms such as The Tech Terms Computer Dictionary

(<https://techterms.com>) name the brevity and frequency of microblogging as its defining features (<https://techterms.com/definition/microblogging>, last accessed March 16th 2017). While posts on Tumblr can be short, they do not necessarily have to be, which suggests that Tumblr might not best be described as a microblogging website. Furthermore, Chang et al. (2014: 21) argue that

Tumblr has more rich content than other microblogging platforms, and it contains hybrid characteristics of social networking, traditional blogosphere, and social media.

Taking into account this existence of “more rich content” (ibid.) and the fact that Tumblr users are not restricted in terms of the length of their posts, a definition of Tumblr as a microblogging website is rejected in this thesis. Since, as implied by the quote above, Tumblr can be considered a hybrid website, it is most likely best understood on its own terms instead of through a comparison to other CMC contexts or websites.

As of March 2017, Tumblr features 340 million blogs and is available in 17 different languages (https://www.tumblr.com/login?redirect_to=%2Fdashboard, accessed March 16, 2017). Due to the fact that registering an account does not require users to provide any personal data besides a valid e-mail address, additional figures are hard to come by and should be taken with a grain of salt. Nevertheless, Golbeck (2015: 152) claims that almost half of Tumblr’s users are aged between 16 and 24 and that the female and male gender are approximately evenly represented with female users accounting for slightly more blogs than male users. However, on Tumblr, where social and political issues like gender identity and expression are discussed and reflected upon by many users, the picture is not as clear-cut as these figures might suggest. My many years of Tumblr experience have taught me that a considerable number of users do not identify themselves as either female or male for example, which Golbeck’s claim about gender distribution seems to completely ignore. Thus, I want to reiterate that sociodemographic data about Tumblr users is not easily obtained and should be met with skepticism. Unless the users provide such sociodemographic information on their blogs, I would argue that they could be considered essentially anonymous (given one does not track their IP-address and thereby locate the device they use to access the Internet). That is not to say that no type of data could be gathered from Tumblr. Data relating to Tumblr’s content and use, for instance, can be obtained, though the sheer amount of blogs means that it

entails a lot of work if one does not limit the sample size. An example study working with Tumblr data is the statistical overview provided by Chang et al. (2014), which provides data gathered from a sample of 586.4 million Tumblr posts. While there can be no guarantee that their figures speak for the entirety of the website, an analysis of a large sample like this does indicate trends on Tumblr, such as the prevalence of photo and text posts (which make up 92% of the posts in their sample; Chang et al. 2014: 22). Overall, Tumblr might be harder to grasp than other websites (such as facebook or youtube, which require sociodemographic information to register) when it comes to sociodemographic features, but this higher level of anonymity should not deter from conducting research related to it.

2.3.2. General layout and functioning

An introduction to Tumblr would not be complete without introducing its general layout and functioning. The present section is thus intended to provide a brief overview. After creating an account by choosing a username and entering a valid email address, a Tumblr user can post texts, videos, images, links, or audio files. All of a user's posts are archived and visible on their blog page. Each blog features a blog page, which is customizable in terms of layout and design and can be reached by entering the username plus '.tumblr.com' into a browser's address bar. The username is thus inextricably linked with the blog's URL. In general, anyone can view a blog if they know the username unless the blog is password protected, in which case the correct password needs to be entered before the blog can be viewed. If a user has created a Tumblr account (and thereby a blog), they can 'follow' other blogs, which means they will see all the blog posts created and shared by those blogs on their dashboard, which acts as a user's 'home base'. The posts appear on the dashboard in reverse chronological order, meaning the most recent post is at the top and older posts are below it. The function of 'reblogging' sets Tumblr apart from many other blogging websites and means that a user shares another user's post on their own blog. Thus, not every post on a Tumblr user's blog has to have been created by this user themselves, but might just have been shared instead. Reblogs can be identified as such because Tumblr automatically adds a reblog sign in the form of two arrows at the top of the post, which shows the blog the post has been reblogged from. This is illustrated in *Figure 3: Reblogged post* below. If a post is not reblogged from the original creator themselves, the post source (aka the original

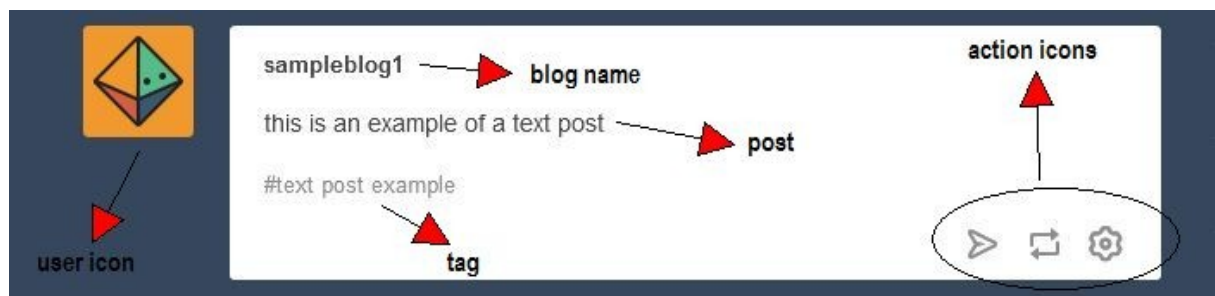
creator) is added below the post. Tumblr also automatically adds the number of times a post has been reblogged in the bottom left corner, under 'notes'. To better illustrate the anatomy of a post as it appears on a user's dashboard, two post examples are depicted below. Figure 1 depicts the most basic layout of a post.

Figure 1: Layout of a post



The different elements of the same post are labeled in Figure 2 below. In general, a post as it appears on the dashboard at the minimum features the user icon, blog name, action icons, tag(s) and of course the content of the post itself.

Figure 2: Layout of a post labeled



If a post is not original, but has been created by another user and reblogged, the layout is somewhat modified, as indicated in Figure 3 below.

Figure 3: Reblogged post



As can be seen in Figure 3, Tumblr automatically adds a reblog icon and indicates not only that the post has been reblogged, but also by whom and from whom. Additionally, as can be seen in the bottom left corner of the post, the website also automatically adds the number of notes, which is the number of times a blog post has been shared.

Figures 1-3 above also show that Tumblr automatically adds buttons to each post which, when clicked, allow the user to perform a number of actions. In these examples, the buttons allow, from left to right, to share the post (that is, send it to another user via the messaging system), to reblog the post, and, since this post was originally created by me, to access the post options. With posts by other users, the latter function is not available, but another function becomes possible: A heart-shaped icon allows users to like the post (which adds the post to a user's 'Likes' folder). All the functions described so far comprise the co-text of a post that is automatically added by the website. What the user adds, then, is the content of the post and the hashtags, or tags, as they are more commonly referred to on Tumblr. In addition to that, users may also add commentary when reblogging a post, which is then added to the post on this user's blog, but not to the post on the original creator's blog. Put simply, if I add the caption *pretty* to a picture posted by another Tumblr user when I reblog said picture, for example, this caption is then visible on my blog, but not on theirs.

2.3.3. Tags on Tumblr

For this thesis, the most relevant feature of posts are tags, since it is their communicative usage that is of interest. In computing, tags are a kind of metadata attached to a file or piece of electronic media, which usually take the form of keywords. They enable the categorization of pieces of media according to the keywords attached to them and are thus a useful tool for search processes within large amounts of data (as is usually the case on websites or the Internet in general). The function and structure of tags on Tumblr is presented in this section in order to provide relevant background information for the analysis. It has already been established that tags can be added to a post by the user. This can be the user who originally posted the post but also every other user who reblogs it. Thus, the same post can appear on Tumblr multiple times with different tags each time. The general function of tags is theoretically like the function of tags on any other website – they are added to a post to make it searchable within Tumblr or within a certain blog and thus provide some organization and structure to the website as a whole and to the specific blog in particular. However, as will be shown in the analysis in chapter 4, this is not the only way tags are used on Tumblr.

Looking at the structure of tags, a number of features are notable. Firstly, it is important to note that tags do not have to be added to a post but are an optional feature. They can thus not be found on every post on Tumblr, since not every user tags and not every user tags every post. Looking back at the layout of a Tumblr post, it becomes clear that tags are the only optional features of a post, which makes investigating them all the more intriguing. Secondly, tags on Tumblr may consist of multiple words, and, contrary to other websites such as Twitter, allow space characters between letters. Therefore, multiple word tags, such as *#such a beautiful sunset* for example are more legible on Tumblr than on websites which do not allow for space characters³ Thirdly, a few limitations apply. According to the Tumblr blog “Unwrapping Tumblr” (www.unwrapping.tumblr.com/) individual tags are limited to 140 characters (<https://unwrapping.tumblr.com/post/121385928727/tumblr-individual-tag-character-limit>, last accessed April 3rd 2017) and 30 tags can be added to each post, but only the first 20 tags are fully functional (<https://unwrapping.tumblr.com/post/121385928727/tumblr-individual-tag-character-limit>, last accessed April 3rd 2017). Thus, tags are not entirely unlimited. Lastly, the

³ The example would then look like this: *#suchabeautifulsunset*

layout of posts on Tumblr causes tags to take a background position since they are added in both a smaller font size and a lighter color. The effect this less pronounced position of tags might have on the way they are used are discussed in the analysis in chapter 4.

2.3.4. Personal Tumblr experience

Before Tumblr is related to some CMC and ELF concepts in the next section, my relation to and experience with Tumblr should be mentioned here. As has already been hinted at above, I have several years of experience with Tumblr. I gathered this experience by being an avid user of the website since June 2011. Consequently, my insights into the website and its workings are based on an extensive amount of time spent on Tumblr, but of course also biased. Therefore, while I will draw on my own experiences from time to time to guide my research and offer additional insights where needed, I will try to keep my analysis as unbiased as possible by asking other Tumblr users for their experiences and observations.

2.4. Tumblr in relation to CMC and ELF research interests and concepts

Since Tumblr is a website, it is situated in a CMC context and can thus be related to several CMC research interests and concepts. Additionally, Tumblr is used by people with various linguacultural backgrounds, many of which decide to communicate on the website via ELF. Therefore, ELF and CMC concepts are equally relevant when considering communication on Tumblr. It is the aim of the present chapter to embed Tumblr in these theoretical contexts and thereby illustrate their relevance for the present research.

Considering that Tumblr is characterized as a type of blogging website, certainly one of the most relevant related research interests is the investigation of blogs and the language use and social parameters associated with them. Blogs, short for 'weblogs' are not easily defined, as boyd (2006) notes, and many bloggers consider trying to define them a fruitless endeavor since it usually means an unnecessarily narrow categorization (boyd 2006: 9). What makes matters worse is that blogs:

are both the product of blogging and the medium through which the blogger produces their expressions. (boyd: 2006. 11)

As a result, they are fluid and changeable, “socially constructed, not technologically defined” (boyd 2006: 11-12) and differ from blogger to blogger, depending on their personal goals, uses, interests and writing styles (Efimova, Hendrick and Anjewierden 2005; Puschmann 2013: 84). This diversity and variability is also reflected on Tumblr and indicated by the plethora of different kinds of content found there. One thing most blogs do seem to have in common, however, is that they are used for “producing digital content with the intention of sharing it asynchronously with a conceptualized audience”, as boyd (2006: 10) puts it. The audience is a conceptualized audience because bloggers do not address the entirety of the general public, but instead imagine an audience that is interested in a blog post’s topic (boyd 2006: 18). As O’Sullivan suggests, new media blur the lines between “large undifferentiated audiences” and “small numbers of familiar interactants” (O’Sullivan 1999) which in turn affects which kinds of relationships are possible within new media contexts. When comparing blogs to other genres, Herring, Scheidt, Bonus and Wright (2004: 2) note that they seem to be a hybrid genre which combines CMC and digital multimedia content. On Tumblr, this multimedia content can take a number of different forms, as already mentioned in the previous chapter. In contrast to Herring et al., however, boyd (2006: 19) argues for an understanding of blogs not as a genre, but as a medium for communication and thus emphasizes the practice of blogs and the fluidity that comes with it. Puschmann (2013: 83) suggests that different understandings of blogs relate to different aspects of blogs that are focused on in the respective research projects but additionally notes that blogs are “increasingly defined in terms of themselves” (2013: 84). However:

What remains unchanged is that blogs structure digital content sequentially and that they are more frequently maintained by individuals than institutions or companies (McLean 2009 as cited in Puschmann 2013: 84).

The question of who maintains a blog is another interesting research interest, since various blogging websites nowadays enable not just professionals, but really anyone to maintain a blog, which results in a growth of user-generated digital content (Puschmann 2013: 87).

One of the main purposes of blogging seems to be self-expression (Puschmann 2013: 88), be it through recounting of stories from a blogger’s personal life or expressing and sharing of personal views (Brake 2012: 1063). Brake (2012: 1066) also mentions “quasi-therapy, quasi-sociality, and blogging as an end in itself”

as prevalent reasons for blogging. Since blogs seem to be rather versatile, it is not surprising that other researchers extend this list of purposes. Such is the case with Qian and Scott (2007) for example, who mention political, educational, promotional, business and marketing uses of blogs. Additionally, blogs embody archives of bloggers' thoughts and interests (Viegas 2005: 3), although this might not always be their main purpose. When it comes to Tumblr, no general claims of purposes can be made, since, as with the content and nature of blogs mentioned above, purposes seem to vary from user to user. I would argue that it is safe to assume, however, that of the purposes mentioned here, each one is represented on Tumblr.

Another interesting aspect of blogs is their level of interaction. Peterson (2011) investigated the conversational nature of blogs and comes to the conclusion that while some blogs might adopt a conversational tone or read like a conversation, this is "not the same as engaging in conversation" (2011: 6, 7). The relation between bloggers and their readers is asymmetrical in terms of expected participation, non-linear and discontinuous (Peterson 2011: 9) and thus, while blogs have the capacity to be somewhat conversational, they do not constitute conversations in Peterson's (2011) view. Comparing this to Tumblr, however, quickly leads to the realization that bloggers on Tumblr can and do indeed interact and engage in conversation, since it is not a pure blogging website but includes social media features (such as a chat function) as well.

Gurak and Antonijevic (2008) investigate the psychology of blogging and pay special attention to the blending of public and private that characterize blogging. They argue that the private aspect of blogs is their content because many bloggers choose to write about themselves and their experience, and the public aspect is the fact that they share these private details with a large audience that in many cases consists of strangers (2008: 64-65). The practice of blogging

incorporates both an old human need—the need for temporal structuring and integrating of past and present experiences—and a new way of doing that—relying on speed, reach, anonymity, and interactivity of Internet communication. (Gurak & Antonijevic 2008: 65)

Related to the concepts of public and private is the concept of identity, which bloggers can construct for themselves on their blogs (Gurak & Antonijevic 2008: 65). Online identity is thus created in the areas of overlap between the public and the private, not distinctly belonging to one or the other. This can also be seen when looking at bloggers' relations to their audience. In a study conducted by Brake

(2012), a number of bloggers were interviewed about their understanding of their respective audiences. While all of their blogs were public, they mostly blogged for themselves and their audiences seemed to not be of great importance (Brake 2012: 1056). In Brake's study, bloggers were aware that there was an audience of people they did not know, and they reported creating an idealized view of their audience, picturing them as interested and sympathetic instead of trying to imagine them as complex human beings with diverse thoughts and behaviors (Brake 2012: 1062). This is also suggested by Viégas, who claims that "bloggers do not feel as if they know their audiences", "have no control over who accesses their entries" and this "leads them to make a number of assumptions" (Viégas 2005: 1). In cases when bloggers reported that they do in fact feel like they know their audience, this usually only comprised the core audience of frequent readers who tended to leave comments (Viégas 2005: 17). Viégas further notes that bloggers knew that this was usually just a small part of their audience, but they nevertheless applied the impression they had of this core audience to the idea of the entirety of their readership (ibid.). Though blogging can bring about interpersonal communication, this connection between bloggers and their audience is not always considered a key aspect of blogging (Brake 2012: 1068), and "interaction in personal blogging can vary" (Brake 2012: 1069). Similarly, Nardi, Schiano and Gumbrecht (2004: 227) note that bloggers in their study valued some interaction with their audience, but overall "desired to keep that audience at arm's length". Nevertheless, wanting to hold one's audience at arms length is not the same as wanting to disengage, as Nardi et al. (2004: 231) suggest. In terms of audience awareness and interaction, Tumblr is a diverse environment. While user may never know the full extent of people a certain post reaches, they are aware of the amount of times it has been shared on Tumblr. Moreover, Tumblr lets its users know how many other users follow their blog through a follower counter. However, this does not mean that bloggers really know their audience and Brake's (2012) and Viégas' (2005) claims are likely to hold true on Tumblr as well. When it comes to interaction, however, Tumblr is likely to feature more of it than is the case in Nardi, Schiano and Gumbrecht's (2004) study, due to its social network functions. It is this higher level of interaction among users that is likely to be the main reason for the need of ELF. The international connections made possible through wider networking need a lingua franca to be realized, and, as Danet and Herring (2007: 22) state, ELF is in a privileged position here. As will be

shown during the analysis in chapter 4, ELF plays a pivotal role when it comes to interaction on Tumblr.

Another aspect related to both blogging practices in general and Tumblr in particular concerns the bloggers' perceptions regarding their own anonymity. These perceptions are influenced by bloggers' target audiences, as Qian and Scott (2007) state. While blogs differ in terms of blogger's self-disclosure and degrees of anonymity, in general, "most bloggers think twice about disclosing highly personal details and private emotions." (Qian & Scott 2007: 1442). In other words, while a higher degree of anonymity might invite bloggers to share more personal details, many bloggers are nevertheless careful about what they do and do not share on their blogs. The concept of anonymity in CMC was a relevant research interest even before blogs first existed. While forms of anonymous communication predate CMC, the technology of computers has certainly made it much easier to remain anonymous (Qian & Scott 2007:1430). Gurak (2001 as cited in Gurak & Anotnoijevic 2008: 63) even goes so far as to consider anonymity as one of the core features of Internet communication, among speed, reach, and interactivity. According to Herring (2002: 135-136), some researchers who subscribe to the cues-filtered-out view of CMC claim that the anonymity that results from the lack of certain cues in CMC has depersonalization and a more egalitarian communicative situation as consequences. The prevalent idea that this determinately results in "boorish behavior" such as rudeness has been criticized (ibid.), however, that is not to say that such behavior does not occur in CMC, as the previously mentioned phenomenon of flaming indicates. When considering anonymity in CMC, it is important to note that anonymity is not a strict dichotomy of anonymous and not anonymous, but can be achieved (or not achieved) to a certain degree, depending on the medium and the type of anonymity (Qian & Scott 2007: 1430). Qian and Scott mention two types of anonymity: visual and discursive anonymity. They explain that visual anonymity entails the physical absence of a message's sender and discursive anonymity is usually understood as the lack of personal information (such as name, gender, or location) given, although it has to be recognized that clues regarding such information can to some extent also be gathered from the way a message is written (ibid.). On Tumblr, both visual and discursive anonymity can be achieved to a high degree, since the only information a user has to offer in order to be able to register for the website is an email address. Moreover, Tumblr also allows for anonymous

messages (if not disabled by the blogger), which means that users can send messages without even revealing their blog name. Whether or not the lack of certain clues is responsible, the fact of the matter is that 'boorish behavior' does take place on Tumblr. However, in line with Herring (2002: 136), it seems unreasonable to assume this is directly caused by the potential for anonymity. While anonymity might encourage disagreeable behavior, it does not cause it, since individual and personal factors are also at play here.

This leads to the subject of identity, which is prevalent in both CMC and ELF research. Androutsopoulos (2006: 423) argues that much of the identity research in CMC takes a social-psychological view and illustrates how identity creation in CMC is enhanced by the option of anonymity. Anonymity offers the affordance of creating different online identities, which may or may not differ from Internet users' identities beyond the boundaries of the Internet (ibid.). Joseph (2004: 1) perceives identity as comprised of two different aspects: a person's name, which serves to differentiate them from others, and an "intangible something that constitutes who one really is". Furthermore, Joseph argues that identity is a linguistic phenomenon and that language and identity are impossible to separate (2004: 12-13). Moreover,

whenever we isolate language from the people who speak and interpret it, we are not getting closer to some kind of essential truth about language. We are getting further from it, toward a generalisation that may well have its uses [...] but can also take the form of a pure abstraction

Both how a person talks and what they talk about influence our perception of a person's identity, which makes clear that identities exist in relation to other people (cf. Joseph 2004: 5). We develop several abstract group identities that connect or separate us from people who do not belong to our groups, and in turn these group identities influence our individual identities. This is also relevant in CMC contexts. If Internet users do not engage with other Internet users, there really is no need for them to create or signal an online identity. If they do wish to use the Internet communicatively, which, as has been mentioned before, usually is the case, they create identities by positioning themselves in a way that has to be interpreted as belonging to certain categories or groups (cf. Joseph 2004: 40). Similarly, Bucholtz and Hall (2009: 18-19) understand identity as emerging discursively in interaction with others. While a subjective 'sense of self' is significant for identity, it can only exist when shared in some form of discourse (ibid.).

According to Bucholtz and Hall:

Identities encompass (a) macrolevel demographic categories; (b) local, ethnographically specific cultural positions; and (c) temporary and interactionally specific stances and participant roles. (2009: 21)

Since identity is contingent on positions and roles, it thus likewise relates to concepts of community. In ELF research, notions of community and identity are central concepts, as indicated in chapter 2.2. As suggested by Šabec (n.d.: 9-11), ELF is not only used to communicate within and signal belonging to an international community of Internet users, but also adopted by non-native speakers of English in order to indicate a certain prestige that comes with it. By using ELF, interactants thus not only widen their reach, but also create an identity they consider favorable and signal community membership. Whether this also holds true for the relation between ELF usage and identity on Tumblr will be shown during the analysis in section 4. When it comes to online identity and relationality, Graham argues that “Identity is dynamic and constantly re-negotiated” and proposes that three aspects are particularly relevant: the perceptions/expectations of audience, the limitations/capabilities of different media, and the goals of interaction (Graham: 2016: 309). What this suggests is that online identities can vary in different CMC contexts if any of these three aspects vary. Like offline identities, they are not entirely rigid entities which necessarily remain unchanged across all CMC situations, but are to be understood as variable and dynamically established. As Barton and Lee (2013: 68) put it:

identity online is not just about *who we are*, but also *who we want to be to others*, and *how others see us*. (emphasis in original)

The Internet makes it easier to actively establish one’s own identity, since others are only partial to the information about oneself that one chooses to disclose. Identity aspects one wants to share with other users are often not only disclosed in some way, but also performed linguistically (Barton and Lee 2013: 68). An example of this is the indication and performance of a desired identity through choosing a certain nickname or screen-name that is considered representative (Barton & Lee 2013: 69). Additionally, performing an identity also takes place through language choice, which again relates to ELF. Tumblr users are not forced to use the website in English, but many of them do so regardless. However, they sometimes switch to their native language when they feel it is called for. Klimpfinger (2007: 39) mentions

two basic reasons for code-switching in ELF interactions: “specifying an addressee” and “introducing another idea” which is considered to be most appropriately expressed in the other language. Both of these behaviors can be observed on Tumblr, especially in relation to national holidays, news, or cultural practices. In these cases, bloggers might feel the need to code-switch in order to signal their local identity and/or culture to other bloggers or to specifically address bloggers who share (part of) that culture or local identity. However, general claims cannot be made and each instance of code-mixing and cultural signaling has to be considered in its discursive context.

So far, online interaction has been discussed with regards to interpersonal interaction and anonymity/impersonal interaction. However, Walther (1996) proposes that there is an additional kind of interaction in CMC, which he refers to as ‘hyperpersonal’ interaction or communication. Hyperpersonal communication is a kind of communication that “has surpassed the level of affection and emotion of parallel FtF [face-to-face] interaction” (Walther 1996: 17). This can be the case if interactants build idealized perceptions of each other (usually based on relatively little information), optimize their self-representation, and benefit from the medium’s characteristics (e.g. editability and convenience) and dynamic feedback processes (Walther 1996: 17-28, Walther 2007: 2539). In hyperpersonal CMC, the interactants benefit from the lack of “interference on environmental reality” (Walther 1996: 33) because they are able to construct and confirm idealized images of themselves and each other (Walther 1996: 29). That way, hyperpersonal CMC can result in more extreme and positive relations than regular face-to-face communication (Walther 2007: 2539). In Walther’s words: “users exploit the technological aspects of CMC in order to enhance the messages” (2007: 2538). It is important to remember, though, that the fact that communication in CMC is mediated does not automatically make it impersonal or hyperpersonal (1996: 33). However, it does provide the opportunity to carry out both of these kinds of communication. Relating the different kinds of interaction to Tumblr it can be observed that all three types of interaction mentioned (interpersonal, impersonal, and hyperpersonal) can be found on the website. Once again, like any of the other aspects discussed before, the types of interactions found vary from user to user and post to post.

All the research interests and concepts considered above – the nature of blogs, relationality between bloggers and readers, anonymity, identity, language

choice, and types of communication – are relevant when trying to characterize Tumblr and Tumblr practices. While I have not been able to find a comprehensive study of Tumblr as a CMC environment and its various characteristics, the section above shows that insights from other fields of research can be applied and used to outline some essential components. Many of the concepts mentioned above will thus come up during the discussion of the research questions and findings in section 4. Moreover, the analysis will bring to light additional connections between CMC; ELF, and Tumblr. Before the research is addressed, however, the next section introduces the research methodology and ethics.

3. Research design

3.1. Research methodology

As has been mentioned already, the present thesis is concerned with the various forms tag usage on Tumblr can take. My general assumption is that many bloggers on Tumblr use tags not just for their intended organizational function, but also make use of them to fulfill communicative needs. This is investigated via a qualitative analysis that aims to provide a more detailed picture of tag usage on Tumblr by answering two research questions:

RQ1: In what ways are tags on Tumblr used to serve communicative purposes?

RQ2: Which ELF key concepts come into play in this type of communication on Tumblr?

The research will take the form of discourse-centered online ethnography (Androutsopoulos 2008) which subsumes some elements of computer-mediated discourse analysis (Herring 2004). These approaches are introduced in what follows.

Discourse-centered online ethnography (henceforth DCOE) is proposed by Androutsopoulos (2008: 2) as an approach that expands a researcher's analysis and interpretation of "what is observable on the screen" by also incorporating online interactants' inputs regarding their linguistic behavior. It thus builds on and goes beyond what Herring labels computer-mediated discourse analysis (henceforth CMDA; Herring 2004) by also incorporating online ethnography of everyday life on the Internet and related input from the language users themselves (Androutsopoulos 2008: 2-4). Since CMDA is in a way the predecessor of DCOE, CMDA is briefly described here before turning to DCOE. Herring suggests CMDA as an approach for

researching online communication and discourse from a linguistic perspective (Herring 2004: 339). CMDA may be supplemented with various other methods, but Herring states “what defines CMDA at its core is the analysis of logs of verbal interaction” (ibid.). What follows from this is that the key method in CMDA is content analysis of log data (Herring 2004: 4). While log data plays an important role in DCOE as well, DCOE also involves additional steps in the research process. Androutsopoulos suggests conducting a systematic observation through online ethnography as a first step in order to be better equipped for data selection, since it gives the researcher an idea of the structure and patterns of the computer-mediated discourse (CMD) they aim to investigate (Androutsopoulos 2008: 5). Not isolated instances of language use are at the center of attention, but processes and patterns which the researcher ideally not only finds in the core but also in the periphery of the field that is the setting for the analysis (Androutsopoulos 2008: 6). Androutsopoulos (ibid.) emphasizes that the researcher must repeat his or her observations, use all available technology and maintain an open mindset during the observation. Once the observation has yielded a sufficient amount of data, this data can be analyzed and, if necessary, used as guidance for additional sampling (Androutsopoulos 2008: 7). The linguistic analysis is then followed by contact with the interactants themselves through interviews which confront them with their own language material (Androutsopoulos 2008: 6). During this step, the interactants are asked to provide an emic perspective and thus add insights that are usually not available through observation and analysis of log data alone (Androutsopoulos: 16). Androutsopoulos (2008: 16-17) argues that this combination of “objective” and “subjective” data is what makes DCOE particularly advantageous compared to other approaches in terms of contextualization and interpretation. For this reason, DCOE was chosen as the approach which guides the methodology for this research project.

The specific research parameters for the present research are as follows. The first step in DCOE, systematic observation, happened over the last two years, during which I closely observed tag usage on Tumblr. However, as has been mentioned before, I have more than 5 years of Tumblr experience at my disposal, which means that I am not just partial to an ethnographic perspective, but also an emic perspective. It is important to note that this might cause a certain bias, but by triangulating these ethnographic and emic perspectives with log data and other

users' insights I hope to achieve sufficient objectivity so as to not taint my research too much.

As is usually the case with DCOE and CMDA, the present research takes a linguistic perspective and aims to analyze a linguistic phenomenon. This is done by analyzing non-randomly selected log data chosen based on its highly illustrative character. Every example of log data analyzed both illustrates the phenomenon of communicative tag usage and was produced by ELF speakers. The selection process was informed by the systematic observation that preceded it. In addition to log data, interview data is gathered from non-randomly selected Tumblr users. These users have been selected because they have been on the website for more than one year, use tags communicatively, use ELF on Tumblr, and were willing to participate in this research project. Due to the fact that I did not interview every single Tumblr user who uses tags, the findings of this research project cannot be generalized or said to apply to tag usage on Tumblr generally, but they are nevertheless indicative of what tag usage by a number of Tumblr users looks like. Additionally, it is not claimed that this phenomenon of communicative tag usage is unique to Tumblr, since, as Hine (2009: 4-5) explains, a clearly defined field site of a particular phenomenon is somewhat suspicious. It may well be that this way of using tags also appears in other CMC contexts that feature tags, but the specific focus in this research is confined to tags on Tumblr. In terms of Kendall's classification of three different boundary choices (Kendall 2009: 22), the present research is limited like this: The basic spatial boundary is the website Tumblr, though not all of the blogs and posts featured there are addressed, so the concrete spatial boundary is a collection of several Tumblr blog posts. The temporal boundary spans over the last two years when it comes to the time during which the analyzed blog posts were written and collected, but the observational period exceeds that boundary by about a year. Lastly, the relational boundary, which refers to the relationship between the researchers and the participants in the study, is not a very strict boundary, since the adoption of DCOE led to direct interaction. Now that the research parameters have been addressed, I turn to important ethical considerations.

3.2. *Ethical considerations*

While the Internet has brought about a plethora of possibilities for researchers, one thing that has to be kept in mind when conducting Internet-related research is the

consideration of ethical issues. That is not to say that ethical considerations are not important in other research as well, but on the Internet, where large amounts of data are available and seem to just wait to be collected, special attention has to be paid to what is and is not ethical to analyze, especially without obtaining permission from the data producers. A key principle in research is the principle of informed consent, which states that research participants need to be given enough preliminary information about the research to give their consent from an informed point of view (cf. Sveningsson 2009: 71). However, if participants are given too much information in advance, research results may be influenced by participants consciously or subconsciously modifying their behavior. Thus, the data gathered might be considered less valid than “real data” gathered without participants knowing about it. Because of this, the Internet has become a popular resource, since it represents a context in which its users provide various kinds of data without being prompted. Nevertheless, not all kinds of data may be obtained without first getting consent, as is stated in the recommendations for ethical decision-making and Internet research devised by Charles Ess and the Association of Internet Researchers (AoIR) ethics working committee (http://www.aoir.org/reports/ethics.pdf_final_version_2002, last accessed March 20th 2017). The committee suggests that obtaining research data in the form of online content without first getting informed consent is tolerable if the data was shared in a public environment (such as public webpages or blogs) and if it does not comprise sensitive material (ibid. 7-8). As Sveningsson notes, this is not as clear-cut a distinction as it seems at first glance, since public and private are not a dichotomy but a continuum, and whether content is seen as sensitive or not can only be assessed by the content creator themselves (2009: 73-83). It is therefore essential for Internet researchers who aim to conduct ethically sound research to consider how these issues are dealt with in their research projects.

The research conducted for this thesis follows the AoIR ethics working committee’s suggestions and obtained informed consent when needed. Technically, with the exception of private blogs, Tumblr is a public platform and Tumblr posts can be viewed by anyone with an Internet connection. Blog posts and their tags thus represent a public data resource. Nevertheless, the blog posts analyzed in this study were anonymized by reformatting them so that only the relevant aspects of the blog posts (a summary of the post and the post’s tags) are depicted and user names or other data are not shared. Additionally, if tag content seemed somewhat sensitive,

bloggers were asked for permission to include and analyze said content in this research project. When it comes to the interview data, all participants were asked to give consent before the data was obtained and had the right to withdraw from the study at any point if they felt the need to do so. Luckily, all of the participants cooperated for the entirety of the study. It is my belief that the research conducted in the next chapter is ethically sound and no participants experienced any harm as a result of it.

4. Analysis and findings

After discussing relevant theoretical background and related concepts as well as the research design, it is now time to turn to the analysis itself. The analysis is split into two parts which discuss one research question each and explore various additional insights gained. Log data analysis makes up the first step and is then followed by and connected to the interview data gathered. The first research question, addressed in chapter 4.1. below, concerns tag usage on Tumblr and inquires about their communicative nature. The second research question, addressed in chapter 4.2., inquires into the ELF nature of communicative tags.

4.1. RQ 1: The communicative nature of tags

4.1.1. The various kinds of tag usage on Tumblr

The observation of tag usage on Tumblr quickly led to the conclusion that tags are definitely not only used for organizational purposes, which the examples analyzed below will showcase⁴. Tags seem to be used for a variety of purposes but judging from the log data the following were considered to be the most prevalent:

- organization and structuring of content
- expressing emotion
- adding commentary (such as opinions or additional ideas)
- relating a post to oneself

I do not claim that these are the only ways tags are used on Tumblr or that this is the ideal way to categorize tag usage, but this is the categorization that emerged out of

⁴ For copyright reasons and anonymization purposes, the examples taken from the log data are presented in a simplified format and their layout thus diverges from the original posts and their tags. The tags themselves, however, have not been altered in any way.

my observational data and is thus the categorization I adopt for the purpose of this research. Furthermore, it is important to note that these ‘categories’ are not always strictly separated and separable from each other but have fuzzy boundaries and can overlap. This is also evident when looking at the examples for each kind of tag usage that are analyzed in the next section. Although the focus of the present study lies on tags when used communicatively, the organizational function of tags is also briefly discussed because an account of tag usage would be incomplete without it.

4.1.1.1. Organization and structuring of content

This first kind of tag usage represents tag usage as it was originally intended, meaning it involves using tags as tools to organize and structure blog posts according to their content. When tagging posts with relevant key words, bloggers make sure that their posts can be found by others who are looking for posts related to those key words. It makes posts searchable and findable among the abundance of posts on Tumblr. Additionally, it enables bloggers to structure their blog content according to relevant topics, themes, or categories and thus helps them sort their content or easily find posts again on their own blog. Consider the following example from the log data provided in Figure 4:

Figure 4: Example 1

description of post content:	a photograph of a ferret sleeping on an a blanket with an otter pattern next to a golden ‘The Niffler’ candle; an image caption which reads: “I ain’t saying he a gold digger...”
tags:	#book #booklr #book worm #book nerd #book pets #book addict #book photography #the niffler #the melting library #book otter blanket #society6 #ferret #diggle bear #basically a niffler #gold digger #bibliophile #harry potter #fantastic beasts #fantastic beasts and where to find them #jk rowling

A couple of things are going on in the tags here. First of all, there are several tags that directly relate to the picture: *#the niffler*, *#ferret* and *#book otter blanket* all name things that are depicted in the picture. On a more abstract level, many of the tags

relate to the type of blog that posted this image (a blog mostly concerned with books and book-related things) and the fact that the blogger wishes for the post to be seen (and, ideally, shared and reblogged) by blogs of a similar nature. Book bloggers are likely to search for tags like *#book*, *#bibliophile*, or *#booklr* (a term many book bloggers use to refer to the community of book blogs on tumblr), or any of the other tags containing the word 'book' and by adding those tags the blogger tries to get their original post to be seen by many people who might be interested in it. The tags *#society6* and *#the melting library* name the online shops where the blanket and the candle can be obtained. Since the depicted candle represents unofficial merchandise for the book and movie "Fantastic beasts and where to find them", a movie within the Harry Potter universe, the tags *#harry potter* *#fantastic beasts* *#fantastic beasts and where to find them* and *#jk rowling* contextualize the post and make it more likely for fans of the movie or book to come across this blog post. The tag *#diggle bear* is only understood with the background knowledge that Diggle is the ferret's name and the blogger generally tags posts with Diggle in it 'diggle bear'. It is unlikely that bloggers who do not know this specifically search for posts tagged 'diggle bear', but the post's creator might use this tag to keep track of their pictures of Diggle. The tag *#gold digger* relates to the image's caption and the fact that the creature called 'Niffler', which the depicted candle is named after, frequently tries to steal gold and valuables. This too, is therefore a tag whose relevance is only fully understood with particular background knowledge. Up until now, all the tags mentioned have an organizational or structuring function, but the blogger also added *#basicall a niffler*, which adds the commentary that the ferret in the picture apparently acts like a Niffler. Thus, this first example already indicates that tags can also be used to communicate information that goes beyond key words used to make the post searchable. Nevertheless, as the analysis shows, most of the tags the blogger added to this post serve an organizational function.

Due to the fact that this example is highly illustrative of organizational tagging and the main research interest lies in more communicative kinds of tags, no further examples of organizational tags are discussed in the following and the analysis from now on concerns itself with more communicative kinds of tag usage.

4.1.1.2. Expressing emotion

Another way tags are frequently used on Tumblr is to express emotion regarding the post that is tagged. This is particularly interesting because one would expect additions of this kind to be added as a reply or commentary to the post, but bloggers frequently express their emotions in and through the tags instead. The possible reasons for this will be addressed in chapter 4.1.2., but for now the analysis is concerned with what emotional tags can look like and will therefore turn to 3 different examples.

Figure 5 below is a rather straightforward example of expressing emotions through tags.

Figure 5: Example 2

description of post content:	picture of two ice-hockey players looking at something on a smartphone and smiling
tags:	#LOOK HOW CUTE #THIS IS SERIOUSLY ADORABLE #THIS MAKES ME SO HAPPY #auston matthews #erik karlsson #toronto maple leafs #ottawa senators

The first notable aspect of the tags for this post is the non-standard use of capitalization. While the first three tags are written exclusively in capital letters, presumably to express excitement or shouting as capitalization often does in CMC (as mentioned in chapter 2.1.), the latter tags do not contain any capital letters at all, even though they represent names. This is not unusual on Tumblr, since capitalization does not make a difference when searching for a tag. Searching for the tag *#apple* will yield the same results as searching for the tag *#Apple*. It is thus not necessary to properly capitalize, and the effort of one keystroke more that it needs to do so can be spared. Nevertheless, as Figure 5 indicates, bloggers still use capitalization, but in many cases they use it to signal paralinguistic features like shouting or (both positive or negative) excitement. Looking at the tags in Figure 5, it can be seen that *#LOOK HOW CUTE*, *#THIS IS SERIOUSLY ADORABLE*, and *#THIS MAKES ME SO HAPPY* are the expression of the blogger's emotion towards the post or picture in the post. They are most likely not intended to serve an organizational function, though that is not to say that one could not go through the

tag *#this is seriously adorable*. However, the search for this tag yields an abundance of very different posts, since what a blogger might consider ‘seriously adorable’ varies from person to person. Tags like *#this makes me so happy* or *#look how cute* thus carry an additional, communicative function, since they communicate something about the blogger to the other bloggers who read them. In this case, they communicate the blogger’s feelings about the post, but possibly also their feelings toward the people depicted in the post. Additionally, the capitalization of the emotion-related tags also hints at the fact the emotions seem to be of a certain strength or potency which the blogger deemed not sufficiently communicated through lowercase letters. Furthermore, this post also indicates that tags for a post do not necessarily have to be of one kind, but that different kinds or functions of tags can be found within the same post. In this case, the first three tags express emotion, while the other four tags refer to what is depicted in the post and are mainly organizational in nature.

As the next example, provided in Figure 6, shows, there are also posts which do not feature organizational tags at all.

Figure 6: Example 3

description of post content:	a set of gifs ⁵ of a scene from the TV show ‘Skam’
tags:	<i>#will this scene ever not make me cry #will i ever watch this scene without tearing up #i’ll give you a hint #answer starts with n and ends with o</i>

Again, like in Figure 5, the tags in Figure 6 still function as organizational tags, but given their communicative and highly conversational nature, they are not likely to be intended as functional, content-structuring tags. Instead, they reveal something about the blogger’s emotion regarding the scene depicted in the post. Reading these tags does not inform other bloggers about the post’s content in any way, but lets them know more about the person who shared them (and the post). For example, the tags *#will this scene ever not make me cry* and *#will i ever watch this scene without tearing up* inform the reader about the fact that the blogger apparently always cries when watching the scene in question. An intention of fulfilling an organizational purpose with tags like this is dubious. Another parallel to the previous example is the

⁵ an image, often animated, in Graphics Interchange Format

aspect of orthographic correctness, or lack thereof, in some cases. Since, as mentioned above, it is not necessary to correctly capitalize in tags, occurrences of ‘I’ spelled ‘i’, such as in the tag *#i’ll give you a hint* in Figure 6, are not infrequent.

Another frequent feature of language use in tags is the inclusion of fixed expressions, as Figure 7 illustrates.

Figure 7: Example 4

description of post content:	a set of gifs from the TV show ‘Gilmore Girls’
tags:	#AWWWWW #my heart #Gilmore Girls #glad queue came

The first tag used in this example, #AWWWWW, is a variation of a common linguistic expression for an [ɔ] or [aw] sound. Depending on the context, this sound can mean different things, which in spoken language is indicated through pitch and duration of the sound. In writing, however, language users have to interpret the meaning via the cotext and context. Furthermore, the number of ‘w’s is used to indicate the duration of the sound. It is impossible to know the exact sound a writer of *awww* had in mind when writing it without asking them, but the number of ‘w’s and, again, the use of capitalization or lowercase letters give some indication of the sound. When considering the many possible variations of spelling and meaning of *awww*, it becomes clear that the tag is not very useful as an organizational tag. It is, however, a useful tag for expressing some kind of emotion. *Awww* is a sentiment that is hard to put into words, since its meaning can vary across different uses and contexts. At the same time, it is a convenient expression because it articulates affective states that might not easily be worded otherwise. Its various meanings range from appreciation, adoration and happiness to compassion, pity, or disappointment. Anyone who shares the knowledge what it can mean and how certain posts are likely to be interpreted by the writer of this tag is in a position to gather the writer’s intention from the context. Thus, the concrete meaning of *awww* has to be negotiated in each communicative situation and is understood through discourse. In Figure 7, the tag #AWWWWW is followed by *#my heart* which is a fixed formulaic expression used by many Tumblr users to expresses an affective state of sentimentality or intense emotion. The combination of #AWWWWW and *#my heart* can thus be interpreted as an expression of either very positive or very negative emotions. Given that the

relating post is a positive post about a TV series the blogger is a fan of, it is much more likely to be an expression of positive emotion in this particular case. This has also been confirmed by the blogger when asked about it. Formulaic expressions and their use on Tumblr are discussed in more detail in chapter 4.1.3. The next two tags, *#Gilmore Girls* and *#glad queue came* are of an organizational nature. While *#Gilmore Girls* is easily understood as a tag naming the TV show the post relates to, an understanding of *#glad you came* requires some Tumblr background knowledge. Tumblr allows users to queue posts, which means they do not have to be shared immediately, but can be put in a 'queue' by the blogger. Queued posts are then automatically posted by the website at certain intervals (which the blogger can determine). Many bloggers create a specific tag for queued posts in order to provide readers with the information that these posts have been queued and not posted 'in real time' by the blogger. This is so as to prevent readers from assuming a blogger is online at a certain time based on the fact that posts are shared on their blog at that time. While the queue tag is often just *#queue*, *#q*, or *#queued*, some bloggers get creative with the language and devise a queue tag that is made up of a certain phrase and replaces one of the words (often a word similar-sounding to queue) with *queue*. In Figure 7, the blogger modified the title of a popular song called 'Glad you came' to read '*glad queue came*'. The queue tag is thus a somewhat organizational tag since it separates queued posts from ordinarily shared posts. Figure 7 is therefore another example which contains different uses of tags within the tags for one and the same post.

There are a countless instances in which tags are used to express emotion, and it is often done by making use of formulaic expression typical for Tumblr. Thus, the examples discussed above just show a small glimpse into this kind of tag usage. A closely related and often overlapping use of tags is the expression of opinion, which is discussed in the next section.

4.1.1.3. Adding commentary

Each time a Tumblr user shares a post, they have the option of adding some form of commentary to the post. For this reason it is all the more notable when users decide to not share their commentaries through this function but express them in their tags instead. The bloggers' motives for this are addressed in chapter 4.1.2. The present section takes a look at some examples wherein different kinds of commentary have

been added as tags of a post. Like the division into different categories of tag usage above, here too the listing of different kinds of commentaries is not meant to suggest that strictly distinct kinds of commentaries exist. However, there seem to be different foci or main purposes for commentaries sometimes. The few examples discussed below do not paint a complete picture of what kinds of commentaries exist in tags on Tumblr, but instead depict examples of some of the more common uses of tags for commenting purposes. Furthermore, it is important to note that in a majority of cases observed and gathered as log data for this thesis, different kinds of commentarial tags are added to one and the same post. Thus, just like it has already been mentioned in the analysis of emotion-related tags, tags which add commentary to one and the same post can fulfill a number of different communicative functions.

adding commentary: reply

In some cases, bloggers use tags to reply to other bloggers, as the example in Figure 8 showcases.

Figure 8: Example 5

description of post content:	a blogger's response to another blogger's thoughts about the TV show 'Shadowhunters'
tags:	#such an amazing analysis! #Jane you are a gem to this fandom #and this is completely true #Magnus knows he's known time and time again that even though he can say something it'll get ignored mostly because he's a downworlder #and the shadowhunters don't seem to the type to take direction well #so why should he spend time on something that is a lost cause rather than something that can make him heard

One of the tags in this example, *#Jane you are a gem to this fandom*, directly addresses one of the bloggers who added to the post by name (the name has been changed here for anonymization purposes). The blogger who added this tag is relatively certain that the blogger they compliment will see and read their tags and, hopefully, appreciate the sentiment. The first three tags, *#such an amazing analysis!*, *#Jane you are a gem to this fandom* and *#and this is completely true* can all be

considered replies to the post and the blogger who posted it. Hence, these tags are clearly used communicatively and not for organizational or structural purposes. Over the course of the following tags, the blogger expresses their own opinions and ideas about the matter discussed in the post, which also represents communicative behavior. Another aspect of tags that comes up when looking at the tags in Figure 8 is their degree of editability. In principle, tags can be edited, but only by retyping the whole tag⁶. Thus, when bloggers make small mistakes like typographical errors or missing a word, as is the case in the tag *#and the shadowhunters don't seem to the type to take direction well*, many of them choose not to correct their mistakes because they do not consider it worth the effort to retype the tag, especially in cases where it is rather lengthy. That is not to say that every blogger is always aware of their mistakes and consciously chooses to disregard them as unimportant, but it is not an unusual occurrence.

adding commentary: opinion

Due to the fact that Tumblr is a website where different kinds of content are shared, it is not surprising that users also share their opinions regarding various contents. Opinions can be shared via replies, comments, and completely separate posts, but also through tags and may range from complete agreement to strong disagreement with a post's content. Why users often prefer to voice their opinions in the tags is discussed in chapter 4.1.2. but for now I turn to some examples of such tag usage.

Example 6 in Figure 9 below features a post in which a blogger expresses their frustration with how the showrunners of the BBC series 'Sherlock' treated their fans during season 4 of the show.

Figure 9: Example 6

description of post content:	text post caricaturing the BBC's dismissive treatment of fans of the BBC series 'Sherlock' who read gay subtext into the show
tags:	#literally #THIS #gay subtext #promoting season 4 in the official accounts with romance filled tweets #and then an episode where #nothing makes sense #and the groundbreaking shit #was euros? #bye

⁶ if a blogger does not use any additional software, that is

The person who shared this post and added the tags seen above expresses their agreement with the sentiment shared in the post by adding *#literally* and *#THIS*. The tag *#literally* is self-explanatory and is meant to express that the blogger believes that this is exactly what occurred, even though the post is written as somewhat of a parody of what happened. *#THIS* represents another common tag on Tumblr which is used to fully agree with a post's statement. In this example, even more emphasis is added by putting it into uppercase letters. The following tags, *#gay subtext*, *#promoting season 4 in the official accounts with romance filled tweets*, *#and then an episode where*, *#nothing makes sense*, *#and the groundbreaking shit*, and *#was euros?* provide additional context and commentary and make it clear that the blogger is disappointed with this turn of events. The final tag *#bye* is relatively ambiguous but can be read as the blogger having lost any interest and removing themselves from the situation, as they explained upon request. An interesting additional observation relates to the way the phrases are sequenced into tags. The breaks between tags do not always follow after complete phrases, as the example *#and then an episode where #nothing makes sense* shows. While *#nothing makes sense* can be a meaningful phrase on its own, it is in this case needed to complete the phrase *#and then an episode where*. The fact that there are breaks between phrases at all is most likely a result of the character limit for tags. According to www.unwrapping.tumblr.com (<https://unwrapping.tumblr.com/post/121385928727/tumblr-individual-tag-character-limit>, last accessed April 3rd 2017), tags which are more than 140 characters long get cut off at the end. As a result, Tumblr users who are aware of this often break up sentences and phrases into smaller units to prevent this from happening. Since tags added to a post are always displayed in the same order, this does not result in intelligibility issues. In this example, the breaks between tags are at seemingly random points, since none of the tags are long enough to be in danger of being cut off, but by breaking off at unexpected points, the blogger seems to somehow convey their feeling of distress to the reader. Due to the fact that tags sometimes require breaks, these breaks become part of the properties of tags and enable writers of tags to use them creatively. It would likely have been much more difficult for the tag user in this example to convey their sentiments in a regular commentary, since random breaks in running text are not conventional language use on Tumblr and possibly

would have been understood differently by the readers. The example in Figure 9 indicates that the affordances and limitations of tags on Tumblr can have an influence on the conventions that arise around them, such as breaking phrases into smaller parts.

adding commentary: additional information

It is not unusual that a blogger who reblogs a post has more information regarding its content and wants to share this information somehow. Here, again, if they want to share this information, bloggers have the choice between adding it as a reply, as commentary below the post, or as tags. Adding this information via tags is a common choice and demonstrated in Figure 10 below:

Figure 10: Example 7

description of post content:	a fan fiction for the TV show 'Yuri on Ice'
tags:	#yoi eq au #myfiction #MANY THINGS TO SAY #you pronounce "Rhys" as 'Rice' -> yes that's a deliberate pun because i am a sad person like that #I didn't give the horses coat colours #because i trust my art friends to fill that in for me #this is working on the feels and also ON MY OWN FEELS #but i am not sorry #cry along with me into the sunset #I GUARANTEE YOU – you'll think twice when you see its raining outside again #love ya all (love is pain)

In this example, the first two tags are of an organizational nature. The tag *#myfiction* simply categorizes the post as the blogger's original fiction, but the first tag, *#yoi eq au*, requires some decoding. Its first element, 'yoi' is the initialism used on Tumblr to refer to the TV show 'Yuri on Ice', and most fans of the show collectively use this tag to signify a post's relation to the show. The second and third part, 'eq' and 'au' refer to the kind of fan fiction found in the main post. In fan fiction terms, 'au' stands for alternative universe and means that some aspect of the original piece of media has been changed in the fan fiction. In this case, the modified aspect is the fact that the writer chose an equestrian setting for their story, denoted by the 'eq' in the tag. The subsequent tags provide additional information about the fan fiction, such as how a

name in it is to be pronounced, why the author did not mention the horses' colors, and the fact that the story apparently causes strong emotions. This last piece of information is given via the tag *#this is working on the feels and also ON MY OWN FEELS*, which features common Tumblr fandom jargon by calling strong emotions 'feels'. In several of the tags in this example the blogger addresses their readers directly, whereby it is clearly indicated that the blogger wrote the tags for them and meant them as a form of communication, which again verifies the supposition that tags can be and are used communicatively on Tumblr.

adding commentary: fan fiction

In some cases bloggers utilize tags to not just provide additional information about their fan fiction, as in Figure 10 above, but even write little pieces of fan fiction directly in the tags of a post. Fandom and fan culture are very prominent on Tumblr (cf. Hillman, Procyk, and Neustaedter 2014) and many blogs and blog posts have something to do with fan activities related to various products of pop culture. One of these fan activities is the writing of fan fiction, which has already been mentioned above and, very roughly defined, represents a transformative involvement with a certain product of pop culture or piece of media by adding something to a story, making up new additions to it, or modifying certain aspects of it. Very short fan fictions can sometimes be found as tags for related posts, as is the case in Figure 11 below.

Figure 11: Example 8

description of post content:	text of an imagined short and humorous dialogue between the characters Stiles and Derek from the TV show 'Teen Wolf'
tags:	#sterek #derek would then roll his eyes in disbelief #and stiles would be like 'it's true i swear i'm so sorry babe' #and Derek would be like 'i know u would never do that to me #now spill' #and stiles would sigh and go 'yeah ur right. like i would even want to try to find someone better than u' #and Derek grins

The first tag, *#sterek*, categorizes the post and its tags as relevant for the relationship between Stiles and Derek from the TV show 'Teen Wolf', in the fan community often referred to as '*Sterek*'. Fans of this relationship know that they can find all kinds of content of interest to them if they search for posts tagged with *#sterek*, and it is rather convenient to have one single designation for the two characters together, since it requires less typing. It is not infrequent that fan communities negotiate 'couple names' for romantic couples they are a fan of, and once a name has been adopted by most of the community, any related content is usually tagged with that name. The blogger in this example seems to be aware of this as well, since they add the tag *#sterek* and thereby make it probable that the post is found by other Sterek-fans. The subsequent tags represent a short fan fiction, or what is more likely considered and labeled a 'ficlet' or 'drabble' because it is extremely short. The tags in Figure 11 are used to communicate a story and as a result carry no real organizational value. Another interesting aspect of the tags in this example is that, once again, there is a certain amount of orthographic variation (i/l, u/you and ur/you're). This behavior might have arisen out of the need for efficiency, or simply "felt right", as the author of these tags reported.

4.1.1.4. Relating a post to oneself

The final variety of tag use discussed here is the utilization of tags to relate a post to oneself. Due to the fact that Tumblr users can not only share original posts, but also posts created by other users, information about themselves can be shared through both original and reblogged posts. With an original post it is usually obvious that the user states something about themselves by posting it, but when posts are reblogged, the user's stance towards the reblogged post can be somewhat ambiguous if they do not further clarify it somehow. Thus, users frequently spell out their relation to the post in the tags. In some cases, such as in the example in Figure 12 below, they also provide additional information about how a post relates to them.

Figure 12: Example 9

description of post content:	text that reads: 'Reblog this if you are literally surprised when people find you attractive'
tags:	#literally me #I was at work handing out fliers when a coworker told me that this guy checked me out #my first thought was 'what really?' #followed closely by me thinking 'ew what?'

Relating a post to oneself can be accomplished through numerous different phrases which are often parts of a kind of Tumblr jargon. In the example in Figure 12 the blogger chose to use *#literally me*. While this information alone would be enough for a reader to know that the blogger shares the sentiment of this post, the blogger in this example goes even further and adds some additional information in the form of an example situation from their life. Anyone who reads the post with these tags added to it thus gains a slightly more in-depth understanding of the blogger's reaction to being considered attractive than people who only read the posts without the tags. Seeing tags as a potential source for additional information about a blogger is a common notion among Tumblr users, as the interview data discussed indicates. In many cases, tags are not just passively consumed, but even actively sought out, as the example in Figure 13 below shows.

Figure 13: Example 10

description of post content:	text reading: 'reblog if youre the kind of person that checks other ppl's tags when they reblog ur stuff'
tags:	#I read all of your tags #ha ha ha ha #they give me life #makes me happy to see what people say #I use tags all the time so I will read all of yours

In this example, both the post and its tags refer to tag culture on Tumblr. With the exception of the tag *#ha ha ha ha*, all of the tags added to this post give the reader information about the blogger's relation to tagging. By adding these tags, the blogger offers additional input about why they read others' tags and how they make them feel and thereby enables the readers to learn something about them. Information like that is not for the blogger's own benefit, but instead represents something they wish to

communicate to their readers. Hence, as with the other uses of tags mentioned above, the communicative nature of tags is obvious here.

All in all, the various examples analyzed above clearly indicate that Tumblr users can and do use tags communicatively. However, just because they can use tags to fulfill a range of communicative purposes, it does not mean that bloggers always do so. It is an option, but not a necessity to use tags that way. Having indicated that tags can be used communicatively, the next question to ask is why Tumblr users choose tags as communicative tools despite being provided with other means of communication on Tumblr. This is addressed in the next section.

4.1.2. Bloggers' reasons for such tag usage

Now that it has been established that communicative tag usage occurs, the next point of interest are the reasons behind it. Therefore, the present section discusses the various reasons for such tag uses that were given by the five bloggers interviewed. In general, the interviewed bloggers do not just use the blogging platform to share content, but also add their own insights and knowledge or share something about themselves by reacting to content. These reactions may take the form of emotions, opinions, or the relaying of additional information and comprise what is typically expressed in communicative tags. This is shared because bloggers feel the need to share it and are afforded the opportunity to do so, which makes up the appeal of social networks (cf. Gurak & Antonijevic 2008).

Two general reasons for this need to share were mentioned by the interviewees: Firstly, a blogger might have additional information or input that could be of interest to their audience. This applies to all the kinds of tag usage wherein some kind of commentary is added or related personal experiences are shared. The second reason is a cathartic function of self-expression. Communicative tags can represent very personal expressions of emotion or opinion and are in many cases not chiefly for the audiences' benefit, but for the benefit of the bloggers themselves. As one interviewee states: "They are just things I need to get out somehow, no matter if anyone listens or not". The interesting question now is why bloggers choose to share these things through tags in particular, and not as comments or replies to the post. A number of reasons for using tags to fulfill this sharing function was compiled from the interview data and roughly grouped according to the following three overarching themes:

- aesthetic
- visibility
- interaction

In order to gain a deeper understanding of the bloggers' reasoning, each of the points listed above will now be explored.

According to the interviewees, the aesthetic and layout of a post (and Tumblr overall) play a major role in their decision to put commentary or other communications into the tags. Since commentaries are added below the original post, they make it much longer and add various context to it. However, in many cases, bloggers do not want to tarnish others' posts this way. As a result, they resort to expressing their thoughts, feelings or opinions via tags when they reblog a post. In doing so, they make these additions visible to their followers, but at the same time keep the post itself free of additions. This is particularly common when it comes to artistic content, as one interviewed blogger notes. Keeping artwork comment-free allows it to stand for itself and be interpreted by each user themselves. As has been claimed by Chang et al. (2014), Tumblr predominantly features visual content, especially in form of various images. Taking this into consideration, it might be less surprising that users add their comments in tags in order to preserve a post's aesthetic and keep it visually pleasing. One interviewee remarked that this is done in order to prevent other users from having to read a number of 'empty' comments (such as a simple *agreed*) when encountering a post. If they really are interested in the other bloggers' opinions, they can turn to the tags, but if not, "tags are easy to ignore". In this view, then, tags are seen as a kind of bonus content that other bloggers might or might not (want to) consume. There is one scenario mentioned in the interviews in which aesthetic considerations do not prevent users from adding commentary directly: if the commentary constitutes meaningful input or active participation in a conversation or discussion.

Related to aesthetic considerations is the notion of visibility. It has been established before that tags are somewhat less visible, since they are presented in a smaller font, lighter color, and below the post. Moreover, while commentary is permanently added to the post (unless it is deleted again), tags are just added to the post as it is reblogged to that specific blogger's blog. One user's specific tags for a post are therefore only seen by the people who see the post through them. While, as established in the previous paragraph, some bloggers tend to use tags in favor of

commenting on a post for aesthetic purposes, others enjoy this lower degree of visibility for other reasons. One chief reason, mentioned by multiple interviewees, is the fact that one can 'hide' behind tags in a way. Therefore, opinions which bloggers might want to share in a more covert way can be expressed in the tags. Opinions might want to be shared in a less visible way due to embarrassment, controversial opinion, or fear of negative consequences, and much of this can be avoided if those opinions or comments are only shared in tags. Tags will not be seen by any user who sees the post, but only by those who follow the user who reblogged it. It is thus likely that those who see it are sympathetic towards the blogger and do not mean them any harm, which means controversial or embarrassing additions in tags in most cases do not spark arguments. The same additions would likely provoke arguments if they were added as commentary, however. As a result, communicative tags might be considered as an option for expressing something in one's own, safe, personal bubble. While bloggers do add something to the post by adding any kind of tag that goes beyond the organizational, they do not thereby actively get involved in the discourse.

This leads to the third theme when it comes to reasons for using tags communicatively: interaction. On Tumblr, interaction is possible but not always what the bloggers are after. According to the interviewees, tags represent a space for sharing thoughts where any kind of reply is not expected or necessary. One user for example mentioned that they use tags to covertly speak to a creator of a post. While directly messaging the creator would seem too frank, thoughts on the post can be expressed through the tags and this way only seen by them if they chose to seek them out. Another interviewee mentioned favoring tags over commentaries when they do not want to start a discussion on the post or add commentary that does not necessitate a reply. One blogger mentioned using tags for getting things off their chest and put it like this: "I needed to say it but I don't need it to be read". While such tags still communicate something to the reader, they do not constitute communication in the sense of discourse or conversation. Reading other Tumblr users' tags is done in order to learn more about the person and their opinions or, as some interviewees reported, checking up on them to see if they are doing okay. They are in most cases not addressed but only read. There is one notable exception to this, however. Several interviewees reported checking up on other bloggers if their tags seemed worrisome and indicated that they needed some (emotional) support.

Behaviors like this showcase that communicative tags are not just ‘shouting into the void’ but effectively communicate something to other people. This clearly indicates that tags fulfill many more purposes than they were originally intended for and are much more interesting than their technical function would suggest. According to the interviewed bloggers, using tags like this and for such diverse purposes is something they picked up from seeing other bloggers do it. It seems to be a behavior typical within certain communities of practice on Tumblr, and by adopting this practice, users may indicate that they consider themselves members of such CofP. The prevalence of a concept of community and CofP on Tumblr will come up again at various other points during the analysis.

4.1.3. Formulaic expressions as notable feature of tag usage

One aspect of tag usage on Tumblr that has to some extent already been indicated in the analysis of the examples above is the utilization of formulaic expressions. These formulaic expressions do not just relate to communicative tagging, but also to ELF practices. Formulaic expressions can be very helpful in ELF communicative situations, since they enable interactants to express a shared notion or meaning without having to put it into the ‘correct’ words themselves. This adherence to the least effort principle, which has already been addressed in chapter 2.2.3. above, seems reflected in tag usage on Tumblr as well. While tag users may want to express their emotions, opinions or personal connection to a post by adding this information in the tags, they do not always make the effort to put it into their own words. In these cases, they make use of formulaic expressions, which are shared and understood across certain communities of practice on Tumblr⁷. While a comprehensive and complete list of formulaic expressions is not the goal of this research, a few examples are now briefly addressed in order to give an idea of some forms formulaic expressions on Tumblr can take.

The elephant in the room is the expression featured in the title of this thesis – *‘I can’t even’*. This expression, like many linguistic variations in CMC, aims to emulate face-to-face spoken interaction by making it seem as if the speaker is not able to finish the sentence. It is usually intended to represent extreme positive or negative emotion or a feeling of overwhelmedness preventing the speaker from

⁷ Note, however, that some of these expressions might also be found in other CMC environments since Tumblr users are not confined to just this website alone.

continuing. *'I can't even'* has been around in tags on Tumblr for years and is still a frequently used formulaic expression now, which stands to reason that it fulfils an effective role for the bloggers who use it.

In some cases, as one blogger mentioned during the interview, formulaic expressions originate in memes (a comedic text in the broader sense of the word text; shared through the Internet) and then find their way into tags. One example for this is the tag *beautiful cinnamon roll, too good for this world too pure*, which, according to the website [www.knowyourmeme.com](http://www.knowyourmeme.com/memes/beautiful-cinnamon-roll-too-good-for-this-world-too-pure) (<http://www.knowyourmeme.com/memes/beautiful-cinnamon-roll-too-good-for-this-world-too-pure>, last accessed March 29th 2017) “is the headline for a satirical article published by *The Onion* in early 2014”, which was then appropriated as a phrase to refer to people or characters a user is particularly fond of. Thus, a blogger might post or reblog a picture of their favorite actor or actress and tag it with *#beautiful cinnamon roll too good for this world too pure* or sometimes just *#beautiful cinnamon roll* to indicate their feelings toward this person. Now that the phrase has been appropriated as a formulaic expression on Tumblr, some bloggers use it even without knowing its origin if they understand its notion and it has come to be efficient even beyond its original context.

Not all formulaic tags used on Tumblr need to consist of that many words, however. In fact, some, like *#same* or *#this*, which are used to signal agreement or identification with a post, only consist of one single word. Naturally, the words themselves already have a meaning in regular sentences, but in certain communities of practice on Tumblr *#same* or *#this* fulfill a very specific function, namely one of indicating identification or agreement and bloggers can thus simply use those succinct expressions instead of having to articulate the sentiment in their own words. The same goes for a more curious tag: *#asdsdgfjgk* or a number of variations thereof. In this case, a blogger expresses a notion similar to *'I can't even'*, with the distinction that they seemingly can not even put their emotions into any words at all and just ‘randomly’ type letters on their keyboard. This also explains the many variations of this tag, since *#asdlknvhf* or *#asdfgh* for example would equally be understood within the CofP wherein the tag is used. Over the course of the observation period for this research project, most of the uses of this tag started with the letter a followed first by some of its neighboring letters on the keyboard and then some letters that are further away. However, the exact orthography does not really

seem to matter since the notion gets across, at least to other members of the CofP, with almost any form the expression might take.

Due to the brevity of the last examples, it might not be entirely accurate to label them formulaic expressions, but nevertheless they are prominent key words on Tumblr which have been attributed with a certain notion and are used, like formulaic expressions, somewhat mechanically. All in all, what the existence of such formulaic tags indicates is that tags on Tumblr a) are used communicatively, and b) feature a linguistic phenomenon that is also frequently found in ELF communication. This directly leads to the second research question, which addresses the relation between tags on Tumblr and ELF.

4.2. RQ2: ELF communication in tag usage on Tumblr

4.2.1. ELF on Tumblr

As has already been established, Tumblr is a website that originated in the United States, but by now also exists in different languages. Due to its American origin, it is not surprising that the English language is prevalent on Tumblr, but there is not just one English language variety featured on the website, and a lot of the English featured on the site is in fact ELF language usage. Adopting Seidlhofer's definition of ELF as "*any use of English among speakers of different first languages from whom English is the communicative medium of choice*" (Seidlhofer 2011: 7, emphasis in original) it can quickly be observed that such language usage occurs on Tumblr, since bloggers with various different native languages use the website and communicate with other bloggers from all over the world using English. However, simply stating that there is ELF usage on Tumblr does not entail much enlightenment, so the following sections will analyze ELF usage in tags on Tumblr in more depth by considering the reasons for it, inquiring bloggers' own perceptions and relating it to relevant ELF concepts already introduced in chapter 2.2. The hope is that this provides a better understanding of ELF communication as it occurs within this particular environment.

4.2.2. Bloggers' reasons for using ELF

Tumblr users could theoretically post content in any language they desire and the website itself is even available in 17 different languages as of writing this thesis, yet

English seems to be the predominant language choice. This perception of English-language dominance was shared by all of the interviewees, with some of them also adding they felt that Tumblr was a very US-American centered site. In view of the website's origin, this is not surprising. Since all of the interviewed bloggers are non-native speakers (NNS) of English but primarily blog and communicate in English on Tumblr, they all qualify as ELF users and were thus asked to share their motivations for using ELF in their blogging ventures. Although this thesis mainly focuses on language use in tags, the interviewees' answers can most likely be considered to be equally valid for their other language use on Tumblr as well. Before turning to the discussion of reasons, I want to reiterate that this research adopts an inclusive view of ELF which does not exclude native speakers from the population of possible ELF users. However, NNS perceptions of ELF on Tumblr were anticipated to be more informative when just considering a small sample size as is the case in this small-scale study, which is why NNS of English were chosen as interviewees.

First and foremost, the interviewed bloggers mentioned reach as a reason for choosing ELF on Tumblr. Due to the fact that many Tumblr users are English native speakers (ENS) and the ones who are not are in many cases considered to be at least somewhat competent EFL speakers, the interviewees argued that by using ELF they could technically reach and be understood by the biggest possible audience (when using just one language). As one user puts it:

Tumblr has a diversified mix of users. It's just easier to communicate with them in English. It's used/taught more around the world. This way I can talk to people from other countries and continents. If I waited for everyone to learn my language to talk to me I'd sit here alone.

A variation of the phrase "everyone understands it" occurred in almost every interview and raises the question whether this is factually true or just a common perception among Tumblr users. Regardless of its factuality, the impression that by using ELF bloggers can make themselves understood by almost every Tumblr user seems very prevalent judging from the interview data. This is also in line with the "internationally-oriented purposes" (Vettorel 2014: 13) for blogging in English mentioned in chapter 2.2.4. above.

However, not just this English-language context was suggested as a reason for language choice on Tumblr, but also the type of content shared. Blogs related to certain movies, TV shows, books, or other products of pop culture are numerous on Tumblr, and the fact that most of these cultural texts and pieces of media are realized

through English influences the fan communities surrounding them. As one interviewee states: “tumblr is catered for fandoms - english was a given”. So even though a user might originally prefer to use the website in their native language, they perhaps soon realize that they find the content and interaction they are looking for within CofP which mostly communicate via ELF. Another user stated that they generally use ELF in any kind of online interaction since, for them, the Internet represents a context where they consider ELF as more fitting to use than their L1.

The notion of a community has already been mentioned above and represents another reason for ELF usage that is intertwined with the considerations of reach and context. It is entirely possible to use Tumblr without making use of ELF or even communicating at all. There is, after all, a lot of visual content that users can enjoy without in any way interacting with other users. However, the interviewed bloggers were all involved in some kind of community of practice and chose to actively participate in discourse and other interactions relating to their specific interests. Considering the nature of CofP it is indisputable that the adoption of a common language within them is advantageous for the communities. Since, as has been established in chapter 2.2.4. ELF is the most commonly used lingua franca in CMC interactions, CofP whose communication is conducted through ELF are likely to be the most inclusive in terms of language, as their potential reach includes all ELF speakers. In short, ELF-based CofP on Tumblr have the widest potential reach and judging from the interview responses, bloggers seem to be aware of this and act accordingly.

One last reason for ELF usage on Tumblr that came up during the interviews relates to identity and was put by a user as such:

It may sound stupid, but I feel much cooler and competent when I use English instead of my first language. And some things (especially emotions) I can't really express in my own language but I have no problem saying them in English because it gives me this sort of distance.

This interviewee seems to identify differently, and more positively, when using English instead of their L1. Their choice is not connected to reach or community, but to their own personal preference and relation to expressing emotion. Choosing ELF on Tumblr, then, comes with added benefits that go beyond expected lingua franca functions such as reach or community, at least for this blogger.

In summary, Tumblr users seem to have a number of reasons for choosing ELF on Tumblr and seem to deliberately choose ELF because it is seen as the most

effective language choice for their communicative purposes. It is a language tool used to ensure the possible inclusion of the largest achievable number of community members when only adopting one language, and new members seem to adapt to this established convention of ELF use if they want to join these CofP. However, agreeing on a common language to be used is not the only ELF-related concept reflected in tags on Tumblr, as the next section shows.

4.2.3. ELF key concepts reflected in language use in tags

In Vettorel's words

in ELF research the main criteria to include a feature as relevant are that it is systematically and frequently occurring in the corpus data, and has been produced by speakers of different linguacultures in multilingual settings (Vettorel 2014: 125)

It is the aim of this section to address a number of ELF features, practices and concepts which occurred in the log data and can be considered as relevant according to Vettorel's specification above. Specifically, the ELF concepts and features regarded below could frequently be found in the log data, were produced by a number of different bloggers with diverse linguistic backgrounds, and featured in different contexts and CofP on Tumblr. While they might not be essential, these practices and features do seem at least relevant to ELF communication through tags on Tumblr.

One typical practice or relevant feature in ELF communicative situations has already been discussed in some detail in chapter 4.1.3. above: the use of formulaic expressions (in particular as an indicator of the least effort principle). Not only do these expressions enable their users to communicate with less effort involved, but they also signal community membership. The initialism *OTP* might not mean anything to photography bloggers, but among fandom-centered blogs, most users likely recognize it as standing for *one true pairing*, a romantic couple one likes the idea of and wishes to be together. Being able to understand and use such phrases and expressions means being a more effective communicator in the CofP they are used in. However, CofP do not only use formulaic expressions but also coin them. Such newly coined phrases (and words) can showcase ELF users' creativity and multilingual resources (Vettorel 2014: 162). Not every new creation will be established as a permanent fixture in the vocabulary of a certain CofP, however. As

Bauer (1983: 43) notes, whether a new coinage is successful depends on the status of the person who created and or used it, the need for it in the first place, and ‘society’s stamp of approval’. Thus, while a number of new creations might appear in bloggers’ tags every day, not all of them get picked up by the CofP. Nonetheless, the fact that some formulaic expressions do gain momentum and are then widely used within one or more CofP still holds true and represents a practice not uncommon in ELF contexts.

Another relevant aspect in ELF communication is the differentiation between ELF and a notion of ‘standard English’ as used by native speakers. As Seidlhofer (2007: 325-326) notes, native speaker conventions are not the be-all and end-all of ELF communication. Similarly, Jenkins, Cogo, and Dewey emphasize that “an ELF perspective sees non-native Englishes as different rather than deficient” (Jenkins, Cogo and Dewey 2011: 283-284). The interviewed Tumblr users seem to agree with this. When asked whether they consider the English they use and read on Tumblr to be different from ‘standard English’, most of them not only agreed, but strongly so. Variations in formality, grammar, and spelling were mentioned as some examples and seem to be in line with some ‘typical’ CMC features mentioned in chapter 2.1.2.2. Some such deviances are illustrated in the example in Figure 14 below:

Figure 14: Example 11

description of post content:	screenshot of a pair of dressage boots in an online shopping cart ; image caption: “oops I did the thing”
tags:	#awks if English sizes are weird and he’s not a medium #and by awks i mean i will ball my eyes out

The tags in Figure 14 feature deviant spelling (*i* instead of *I*), the shortening of *awkward* to *awks* and the incorrect word choice of *ball* instead of *bawl* in the phrase ‘to bawl one’s eyes out’ if one were to go by ‘standard English’ norms. However, none of these deviations were criticized or even pointed out by other bloggers and the intended meaning of the tags can still easily be discerned. Thus, adhering to standard norms does not seem to be considered vital in this situation. In addition to deviant spelling, grammar and lower degrees of formality, some interviewees also mentioned a kind of Tumblr slang, which consists of formulaic expressions, initialisms

(such as OTP or TBR⁸ for example) and deviant primary meanings of regular words such as 'ship'. When there is talk of a ship on Tumblr it is much more likely to mean a romantic relationship than a kind of boat. These expressions, initialisms and deviant meanings have to be learned by a user if they want to be part of the CofP that use them. The notion of various CofP is particularly important here since Tumblr is a vast website with numerous users who all use language differently, which negates the notion of a kind of 'Tumblr language'. As one interviewee notes, writing styles and language use are very variable from blogger to blogger. While some users might adapt to certain CofP and make use of expressions established therein, others might adhere to their notion of 'standard English'. However, users usually do not seem to demand impeccable language use or correct other users' mistakes. Thus, the let-it-pass principle (Firth 1996) seems to apply in this ELF context, especially when it comes to (the less visible) tags. In all of the log data, there was not a single instance in which a blogger acted on the need to correct another blogger's tags in terms of grammar, spelling, or style. Being able to understand each other seems to be the prime concern. That being said, some interviewees did report that they pay attention to their own writing and take care to adhere to 'standard English' norms because they felt insecure about their English competence, despite not expecting any negative consequences from norm-deviating language use. In this case it seems to be a matter of wanting to present oneself in the best possible light. Nonetheless, judging from the log data the let-it-pass principle still holds true when it comes to interactions with other users.

Although some users adhere to a notion of 'standard English' and some argue they generally use ELF with Tumblr-specific expressions and phrases, there are also instances in which users signal their personal linguacultural backgrounds through practices like code-mixing. While, as is typical in ELF interactions, the communicative goal takes precedence, tags sometimes also feature non-English words or phrases which relate to the post. Such is the case in tags for posts that address very local happenings or subjects for example. One interviewed blogger mentioned posting about country-specific holidays and news. Figure 15 shows the tags for one of their posts about a typical Romanian tradition:

⁸ 'to be read', usually referring to a 'to be read' list or collection of books

Figure 15: Example 12

description of post content:	an image of a book and a red-and-white bracelet with small charms; an image description introducing the tradition of giving a măștișor (charm or talisman) on the 1 st of March
tags:	#mary rambles #martisor #Romanian traditions #gone girl #book photo #mine

Since the post is all about the Romanian tradition of offering people a talisman called măștișor on the 1st of March, the blogger considered it more meaningful to tag the post with *#martisor* instead of *#charm* or *#talisman*. The subsequent tag *#Romanian traditions* informs the reader of the cultural context of the post, but also hints at the fact that the blogger who posted it obviously has knowledge regarding this culture. Another instance of cultural signaling in tags relates to the cultural context of a specific piece of media, not the blogger themselves. This is illustrated in Figure 16 below, which depicts the tags for a post related to the Norwegian TV show ‘Skam’, which, as of writing this thesis, has not officially been translated into English but still reaches an international audience through unauthorized translations provided by ELF speakers who are Norwegian native speakers.

Figure 16: Example 13

description of post content:	a picture of two pieces of fan art for the show ‘Skam’
tags:	#he takes care of his love #softest boyfriends alive #du er ikke alene #otp: kardemomme #skam art #fanart #isak x even #isak valtersen #even bech næsheim #skam #skam spoilers #skam sesong 3

In this example, three instances of code-mixing (aside from the name of the show, which in itself is a Norwegian word meaning ‘shame’) can be found. The tag *#du er ikke alene* (‘you are not alone’) quotes a catchphrase related to the then current season of the show, acknowledged through the tag *#skam sesong 3* (‘skam season 3’). The third Norwegian utterance is found in the tag *#otp: kardemomme* (‘otp: Cardamom’), which categorizes the couple in the post as a ‘one true pairing’, and

uses a reference from the show – *kardemomme* – as a key word to refer to them. All three of those tags represent code-mixing and a use of the show's original language, Norwegian, and are not only understood by Norwegian native speakers, but also by international fans of the show who have picked up these phrases when joining the 'Skam' fan community. It thus becomes apparent in this example that cultural signaling in tags on Tumblr does not always need to be cultural signaling of the blogger's own culture, but may in fact be cultural signaling related to a TV show, book, movie, or other piece of media. Judging from this and similar examples, it can be argued that Tumblr is not just an environment where different people signal and share their own culture, but also a setting where users can learn about and even adopt aspects of other cultures if they feel inclined to do so. It is precisely in behaviors like this that the international or even global nature of the website comes into view. Further research into cultural behavior on Tumblr and how it relates to other facets of the website would likely yield interesting insights.

All in all, a number of features that have been considered relevant in other ELF-related studies seem to be of relevance when it comes to communication on Tumblr as well. While much of the previous research has focused on ELF communication in business and educational domains (Cogo et al. 2011: 285), other domains such as tourism, politics, technology and the media are increasingly studied nowadays (Cogo et al. 2011: 297). Due to the fact that different concepts and ELF features were derived from observations and analyses of communication in different domains, they might not hold true across every ELF situation. Blanket statements about ELF communications thus seem to hold little value, which is why the present section aimed at reviewing key concepts and practices such as making use of formulaic expressions, signaling cultural identity, and giving precedence to communicative success over adhesion to standard norms in relation to tag use on Tumblr. Thereby, ELF-related insights into the domain of popular culture, media and its discussion via CMC have been gained. In this context, the log and interview data shows that several claims about ELF practices and relevant features in business or educational domains have held true when it comes to the domain of popular culture and media as presented in fan practices on Tumblr as well. ELF communication in tags on Tumblr seems to be content-oriented instead of defined by standard norms, negotiated within and shaped by different communities of practice, and nevertheless still influenced by individual identities and linguacultural backgrounds. The final

product might deviate from bloggers' notions of 'standard English', but as long as successful communication is achieved, this is not considered problematic. Judging from the analysis and bloggers' own inputs regarding their ELF usage, the ELF communications in tags on Tumblr observed and addressed in this research project seem to obtain many similarities with ELF communicative situations investigated in other studies. Several ELF practices and concepts proposed elsewhere, such as the least effort principle, the let-it-pass principle, and signaling of cultural identity feature in and apply to the data for this study as well. It seems safe to assume that they are thus relevant in terms of Vettorel's consideration stated at the beginning of this chapter (Vettorel 2014: 125) and further study of ELF usage on Tumblr might prove insightful.

5. Conclusion and research perspectives

This thesis set out to provide a detailed picture of tags as they are used as communicative tools on Tumblr by means of a qualitative analysis in the style of discourse-centered online ethnography (Androutsopoulos 2008). By triangulating ethnographic observation of the phenomenon, log data gathered from the website, and interview data gathered from a small number of Tumblr users, two research questions were addressed and answered through a qualitative analysis.

Research question one inquired into the communicative use of tags and the analysis showed that there are a number of ways in which tags are used communicatively. Any tag that goes beyond a keyword related to a post's content and intended to somehow organize or structure said content is likely to fulfill a kind of communicative function. Tumblr bloggers utilize tags, a basic function of the website, in new and sometimes even creative ways and thereby showcase that elements of new technology can be turned into communicative tools, even if they were not originally intended to be used that way. From expressing emotion to adding different kinds of commentary or relating a post to oneself, the examples of Tumblr posts and their tags discussed in the analysis displayed a variety of tag uses and, based on examples like *#this makes me so happy* or *#and this is completely true*, it can safely be concluded that tags as they are used in some posts on Tumblr surpass a purely organizational or structural function. A discussion of examples of each tag use encountered in the observation provided a glimpse at the varied character of

communicative tags and highlighted aspects such as the occurrence of deviant orthography common in many CMC contexts (e.g. capitalization for emphasis such as in *#LOOK HOW CUTE*, repetition of letters for indicating pronunciation and/or tone such as in *#AWWWWW*, and lexical creativity as exemplified in the word ‘*feel/s*’), the different sentiments expressed in tags (e.g. fondness, exasperation, anger, curiosity, or emotionality in general), and the use of formulaic expressions or newly-coined words as markers of group memberships. Tags were shown to be not only highly variable from blogger to blogger and context to context, but also evidently complex in their functions. Not only do they structure content, but they are also used to communicate information about the bloggers themselves, and can, but do not have to be decidedly personal in nature. Bloggers’ reasons for such tag usage are as diverse as the form their tags take, but the interviewed bloggers favored tags over commentaries as the site for their communicative utterances mainly due to factors related to aesthetic considerations, visibility, and interactional nature of tags. Put in a nutshell, Tumblr users seem to enjoy the fact that tags do not tarnish a post, are less visible than commentaries, and do not demand interaction while still being communicative. Using tags as a ‘secret whisper space’ is a common notion on Tumblr. One notable feature of tag usage addressed in the analysis was the establishment and use of formulaic expressions within different communities of practice. Tumblr users establish different formulaic expressions and some of these expressions then gain momentum and are commonly used within one or several CofP. The expression *I can’t even*, which is part of the title for this thesis, is one of such expressions and is valued by the bloggers because they feel that it expresses a certain sentiment so well that it is best put in those exact words. Other users who have shared this sentiment before and are aware of the expression’s denotation because they belong to the CofP it is used in then know exactly which feelings the user wants to articulate by stating *I can’t even*. As the interviewed bloggers reported, using such shared expressions and other group practices play a role in shaping and signaling a certain group identity. While group membership is certainly not necessary for using the website, many Tumblr users, among them all of the interviewees, do prefer to join and partake in CofP. One relevant aspect of any kind of community is the need for a shared means of communication, which leads to research question number two.

The second research question inquired about the ELF nature of communicative tags and the various key concepts related to it. Seeing as Tumblr originated in the United States but has spread and is now used by bloggers all over the world, it is not surprising that English and ELF play a big role when it comes to language use on the website. Tumblr users seem to consider ELF as the most beneficial language choice on Tumblr. An analysis of the interview data gathered from bloggers who are all non-native speakers of English brought to light the following reasons for choosing and using ELF: reaching the biggest possible and diverse audience of other bloggers, adapting the language to the language of the piece of media that is discussed, participating in a community of practice, and expressing identity. While not every Tumblr user communicates with the website's diverse user base, the log data indicated that many users do and can as a result be considered ELF speakers. In consequence, it is conceivable that prevalent ELF concepts and practices are reflected in ELF communications on Tumblr. The existence and prevalence of ELF practices and central concepts hypothesized in other contexts was indeed demonstrated in Tumblr communications through the analysis of both log data and interview data. The occurrence of formulaic expressions, dissociation from 'standard English' norms, signaling of cultural identity or notions of communities of practice, which have all been introduced as relevant to ELF in different communicative situations, seem to prominently feature in the domain of social media and blogging, in particular on Tumblr, as well. However, this is only indicated and the present research does not serve as basis for any generalizations about ELF use in CMC or on Tumblr. Nevertheless, this research provides qualitative insights into the forms ELF usage in tags on Tumblr can take and thus presents a contextualized view of ELF as it is used in this medium.

Overall, it is my hope that this thesis contributes to the fields of both CMC and ELF in relation to the particular context of communicative tags on Tumblr and provides relevant new insights to Internet linguistics. It provides an informed understanding of the phenomenon of communicative tag use on Tumblr and the reasons behind it, a contextualized view of ELF as it is used in such tags, and additional insights gathered not only through observation and log data, but also through interviewing Tumblr users about their observations. On the whole, the analysis revealed that tag usage is highly varied in terms of purpose and the form it

takes, but there seems to be agreement among the participants regarding the choice of ELF as the preferred means of communicating via tags.

The research project undertaken in this thesis barely scratches the surface of CMC and ELF research respectively and with the Internet's rapid growth, new linguistic phenomena and consequently new research interests arise continuously. Various affordances of different CMC contexts generate opportunities for new linguistic behaviors and tremendous amounts of data. However, it is important to be aware of ethical considerations and note that even though such data is easily obtainable online, that does not automatically justify gathering it without consent. Within this research project, I have adhered to the guidelines provided by the Association of Internet Researchers (AoIR) ethics working committee (<http://www.aoir.org/reports/ethics.pdf> final version 2002, last accessed March 20th 2017) and would thus argue that the research has been conducted in good conscience. It is my belief that Internet research benefits most when Internet users do not feel invaded in any way and any research within CMC contexts should therefore be conducted with ethical considerations in mind.

There are many Internet- and ELF-related research directions which have not yet been explored in depth or at all. One of these is the study of written ELF communication, particularly on the internet, which so far pales in comparison to the studies conducted around spoken ELF interactions. Since the Internet provides a major channel for international communication, many modes of communication taking a written form, there are bound to be interesting ELF-related insights to be found. Another flourishing and seemingly ever-growing area for Internet research relates to the various interactions and other linguistic behaviors found in CMC contexts, which change and develop frequently. The World Wide Web in particular, with large numbers of blogs, social media websites or other sites for interaction represents a gold mine of data which could lead to many a new research interest. However, as mentioned above, ethical considerations have to be kept in mind. Due to the fact that the Internet and the World Wide Web within it continuously grow in importance for our daily lives, insights into linguistic behavior in CMC contexts might also be of relevance when it comes to language teaching. However, due to time and space constraints, this is not further discussed here. All in all there is a plethora of opportunities for further Internet- and ELF-related research within CMC as a whole, but also within specific limited contexts such as the website Tumblr. While it cannot

be denied that some linguistic phenomena on the Internet can be very short-lived, there is also a growing body of as of yet uninvestigated phenomena which seem to be more permanent. This relative permanence is also the case for the phenomenon of communicative tags as explored in this thesis. Therefore, I hope to have made a relevant contribution to CMC and ELF research and shown that websites such as Tumblr are worth investigating when it comes to linguistics, although I do not doubt that cultural studies would find value in the website as well, and am curiously awaiting the future directions research into various (linguistic) behaviors on Tumblr might take.

References

- Androutsopoulos, Jannis. 2006. "Introduction: Sociolinguistics and computer-mediated communication". *Journal of Sociolinguistics* 10(4), 19–38.
- Androutsopoulos, Jannis. 2008. "Potentials and limitations of discourse - centered online ethnography". *Language@Internet* 5, article 8. Available at: <http://www.languageatinternet.org/articles/2008/1610> (last accessed on February 24, 2017).
- Androutsopoulos, Jannis. 2010. "Localizing the Global on the Participatory Web". In Coupland, Nikolaus (ed.). *The handbook of language and globalization*. Oxford: Wiley-Blackwell, 203-231.
- Androutsopoulos, Jannis. 2013. "Code-switching in computer-mediated communication". In Herring, Susan; Stein, Dieter; Virtanen, Tuija (eds.). *Handbook of the pragmatics of CMC*. Berlin: Mouton de Gruyter, 667-694.
- Angouri, Jo. 2016: "Online communities and communities of practice". In Georgakopoulou, A.; Spilioto, T. (eds.). *The Routledge Handbook of Language and Digital Communication*. New York: Routledge, 323-338.
- Baker, Joseph. 2009. "The cultures of English as a Lingua Franca". *TESOL Quarterly* 43(4), 567-592.
- Baker, Will. 2015. *Culture and identity through English as a lingua franca: Rethinking Concepts and Goals in Intercultural Communication*. Berlin: Walter De Gruyter.
- Baron, Naomi. 2003. "Language of the internet". In Farghali, Ali. (ed.). *The Stanford Handbook for Language Engineers*. Stanford: CSLI Publications, 59-127.
- Barton, David; Lee, Carmen. 2013. *Language online: Investigating Digital Texts and Practices*. New York, NY: Routledge.
- Bauer, Laurie. 1983. *English word formation*. Cambridge: Cambridge University Press.
- Baym, N. K. 2000. *Tune in, log on. Soaps, Fandom, and on-line community*. Thousand Oaks, CA: Sage.
- Bieswanger, Markus. 2013. "Micro-linguistics structural features of computer-mediated communication". In Herring, Susan C.; Stein, Dieter; Virtanen, Tuja. (eds.). *Pragmatics of computer-mediated communication*. Berlin: de Gruyter, 463-485.
- Block, David. 2004. "Globalization, Transnational Communication and the Internet". *International Journal on Multicultural Sciences* 6(1), 22-37.

- Bourlai, Elli; Herring, Susan. 2014. "Multimodal communication on tumblr: 'i have so many feels!'". Proceedings of the 2014 ACM conference on web science, 23 June 2014, 171-175.
- boyd, danah. m. 2006. "A blogger's blog: exploring the definition of a medium". *Reconstruction* 6(4). <http://www.danah.org/papers/ABloggersBlog.pdf> (last accessed 23 February 2017).
- Brake, David. 2012. "Who Do They Think They're Talking To? Framings of the Audience by Social Media Users". *International Journal of Communication* 6, 1056-1076.
- Bucholtz, Mary; Hall, Kira. 2009. "Locating Identity in Language". In Llamas, Carmen (ed.). *Language and identities*. Edinburgh: Edinburgh University Press, 18-28.
- Chang, Yi; Tang, Lei; Inagaki, Yoshiyuki; Liu, Yan. 2014. "What is Tumblr: A Statistical Overview and Comparison". ACM SIGKDD Explorations Newsletter, 25 September 2014, Vol.16(1), 21-29.
- Cogo, Alessia. 2012. "English as a Lingua Franca: concepts, use, and implications". *ELT Journal* 66(1), 97-105.
- Crystal, David. 2003. *English as a Global Language*. Cambridge: Cambridge University Press.
- Crystal, David. 2005. "The scope of Internet linguistics". American Association for the Advancement of Science meeting, 18 February 2005.
- Crystal, David. 2006. *Language and the internet*. (2nd edition). Cambridge: Cambridge University Press.
- Crystal, David. 2011. *Internet linguistics: A student guide*. London: Routledge.
- Danet, Brenda; Herring, Susan C. 2007. "Introduction: Welcome to the Multilingual Internet". In Danet, Brenda; Herring, Susan C. (eds.). *The Multilingual Internet: Language, Culture, and Communication Online*. Oxford: Oxford University Press, 3-39.
- Dürscheid, Christa. 2004. „Netzsprache: ein neuer Mythos“. In Beißwenger, Michael; Hoffmann, Ludger; Storrer, Angelika. (eds.). *Internetbasierte Kommunikation*. Oldenburg: Red. OBST, 141-157.
- Efimova, Lilia; Hendrick, Stephanie; Anjewierden, Anjo. 2005. "Finding 'the life between buildings': An approach for defining a weblog community". In: *Proceedings of Internet Research 6.0: Internet Generations*.
- Ferguson, Gibson. 2012. "The practice of ELF". *Journal of English as a Lingua Franca* 1(1), 177-180.
- Firth, Alan. 1996. "The discursive accomplishment of normality. On 'lingua franca' English and conversation analysis". *Journal of Pragmatics* 26(2), 237-259.

- Firth, Alan. 2009. "The lingua franca factor". *Intercultural Pragmatics* 6(2), 147-170.
- Golbeck, Jennifer. 2015. *Introduction to Social Media Investigation: A hands-on approach*. Waltham, MA: Syngress.
- Graham, Sage Lambert. 2016. "Relationality, friendship, and identity in digital communication" In Georgakopoulou, A.; Spilioto, T. (eds.). *The Routledge Handbook of Language and Digital Communication*. New York: Routledge, 305-320.
- Gurak, Laura J.; Antonijevic, Smiljana. 2008. "The psychology of blogging: You, me, and everyone in between". *American Behavioral Scientist* 52(1): 60–68.
- Herring, Susan C. 2001. "Computer-mediated Discourse". In Schiffrin, D.; Tannen, D.; Hamilton, H. (eds.). *The Handbook of Discourse Analysis*. Oxford: Blackwell, 612-34.
- Herring, Susan C. 2002. "Computer-mediated communication on the Internet". In Cronin, B. (ed.). *The Annual Review of Information Science and Technology*. Medford, NJ: Information Today Inc./American Society for Information Science and Technology, 109-168.
- Herring, Susan C. 2004. "Computer-Mediated Discourse Analysis: An Approach to Researching Online Behavior". In Barab, Sasha; Kling, Rob; Gray, James (eds.). *Designing for Virtual Communities in the Service of Learning*. New York: Cambridge University Press, 338-376.
- Herring, Susan C.; Scheidt, Lois Ann; Bonus, Sabrina; Wright, Elijah 2004. "Bridging the gap: A genre analysis of weblogs". Proceedings of the 37th Hawai'i International Conference on System Sciences (HICSS-37). Los Alamitos: IEEE Computer Society Press.
- Herring, Susan C.; Stein, Dieter; Virtanen, Tuja. 2013. "Introduction to the pragmatics of computer-mediated communication". In Herring, Susan C.; Stein, Dieter; Virtanen, Tuja. (eds.). *Pragmatics of computer-mediated communication*. Berlin: de Gruyter, 3-32.
- Hillman, Serena ; Procyk, Jason ; Neustaedter, Carman. 2014. "'alksjdf;lksfd': Tumblr and the Fandom User Experience". Proceedings of the 2014 conference on designing interactive systems, 21 June 2014, 775-784.
- Hine, Christine. 2009. "How Can Qualitative Internet Researchers Define the Boundaries of Their Projects?" In Baym, Nancy K.; Markham, Annette N. (eds.). *Internet Inquiry*. Thousand Oaks: SAGE Publications, 1-20.
- Hülmbauer, Cornelia. 2013. "From within and without: the virtual and the plurilingual in ELF". *Journal of English as a Lingua Franca* 2(1), 47-73.
- Jenkins, Jennifer; Cogo, Alessia; Dewey, Martin. 2011. "Review of developments in research into English as a lingua franca". *Language Teaching* 44(3), 281-315.

- Joseph, John. 2004. *Language and Identity*. Basingstoke: Palgrave Macmillan.
- Kendall, Lori. 2009. "A Response to Christine Hine" In Baym, Nancy K.; Markham, Annette N. (eds.). *Internet Inquiry*. Thousand Oaks: SAGE Publications, 21-25.
- Klimpfinger, Theresa. 2007. "Mind you, sometimes you have to mix'—The role of code-switching in English as a lingua franca". *Vienna English Working Papers(VIEWZ)* 16(2), 36-61.
- Kramsch, Claire J. 1998. *Language and culture*. Oxford: Oxford University Press.
- Leppänen, Sirpa; Pitkänen-Huhta, Anne; Piirainen-Marsh, Arja; Nikula, Tarija; Peuronen, Sarija. 2009. "Young People's Translocal New Media Uses. A Multiperspective Analysis of Code Choice and Heteroglossia". *Journal of Computer-Mediated Communication* 14, 1080-1107.
- Maiz-Arevalo, C; Garcia-Gomez, A. 2013. "'You look terrific!' Social evaluation and relationships in online compliments". *Discourse Studies* 15(6), 735-760.
- Mauranen, Anna. 2003. "The corpus of English as a lingua franca in academic settings". *TESOL Quarterly* 37(3), 513–527.
- McLean, Jennifer. 2009. "State of the blogosphere 2009". Technorati.com, October 19.
- Meierkord, Christiane. 2002. "'Language stripped bare' or 'linguistic masala'? Culture in lingua franca conversation". In Knapp, Karlfried; Meierkord, Christiane (eds.). *Lingua franca communication*. Frankfurt am Main: Peter Lang, 109-133.
- Muniandy, Alageswary V. A. (2002). Electronic discourse (E-discourse): Spoken, written or a new hybrid?. *PROSPECT-ADELAIDE-*, 17(3), 45-68.
- Murray, Denise. 2000. "Protean Communication: The Language of Computer-Mediated Communication". *TESOL Quarterly* 34(3), 397-421.
- Murray, Peter J. 1997. "A rose by any other name". CMC Magazine. January 1997, <http://www.december.com/cmc/mag/1997/jan/murray.html> (last accessed April 21st 2017)
- Nardi, B. A.; Schiano, D. J.; Gumbrecht, M. 2004. "Blogging as a social activity, or would you let 900 million people read your diary?". Proceedings of the 2004 ACM conference on computer supported cooperative work. New York: ACM, 222-231.
- Norton, Bonny. 2010. "Language and Identity". In Hornberger, Nancy; McKay, Sandra Lee. *Sociolinguistics and Language Education*. Bristol: Multilingual Matters, 349-369
- O'Sullivan, P. 1999. "Personal Broadcasting: Theoretical Implications of the Web", <http://www.ilstu.edu/~posull/PersBroad.htm> (last accessed February 10, 2017).

- Page, Ruth; Barton, David; Unger, Johann; Zappavigna, Michele. 2014. *Researching Language and Social Media. A student guide*. London: Routledge.
- Paolillo, John C. 2007. "How Much Multilingualism?". In Danet, Brenda; Herring, Susan C. (eds.). *The Multilingual Internet: Language, Culture, and Communication Online*. Oxford: Oxford University Press, 408-430.
- Pavlenko, Aneta; Blackledge, Adrian. 2004. "New Theoretical Approaches to the Study of Negotiation of Identity in Multilingual Contexts". In Pavlenko, Aneta; Blackledge, Adrian (eds.). *Negotiation of Identities in Multilingual Contexts*. Clevedon: Multilingual Matters, 1-33.
- Pennycook, Alastair. 2010. "Nationalism, Identity and Popular Culture". In Hornberger, Nancy; McKay, Sandra Lee. *Sociolinguistics and Language Education*. Bristol: Multilingual Matters, 62-86.
- Peterson, Eric E. 2011 "How conversational are weblogs?". *Language@Internet* 8, <http://www.languageatinternet.org/articles/2011/Peterson> (last accessed March 15, 2017).
- Pitzl, Marie-Luise. 2009. "We should not wake up any dogs': Idiom and metaphor in ELF". In Mauraanen, Anna; Ranta, Elina. (eds.) *English as a Lingua Franca. Studies and Findings*. Newcastle upon Tyne: Cambridge Scholars Publishing, 298-322.
- Pölzl, Ulrike; Seidlhofer, Barbara. 2006. "In and on their own terms: the 'habitat factor' in English as a lingua franca interactions". *International Journal of the Sociology of Language* 177, 151-176.
- Posteguillo, Santiago. 2002. "Netlinguistics and English for internet purposes". *Ibérica* 4, 21-38.
- Puschmann, Cornelius. 2013. "Blogging" In Herring, Susan C.; Stein, Dieter; Virtanen, Tuja. (eds.). *Pragmatics of computer-mediated communication*. Berlin: de Gruyter, 83-108.
- Qian, Hua; Scott, Craig R. 2007. "Anonymity and self-disclosure on weblogs". *Journal of Computer-Mediated Communication* 12(4), 1428-1451.
- Risager, Karen. 2006. *Language and culture: Global flows and local complexity*. Clevedon, England: Multilingual Matters.
- Šabec, Nada. n.d. "Slovene-English netspeak: linguistic and socio-cultural aspects". http://oddelki.ff.uni-mb.si/filozofija/files/Festschrift/Dunjas_festschrift/sabec.pdf (last accessed April 1, 2017).
- Schmidt, Jan. 2007. "Blogging practices: An analytical framework". *Journal of Computer-Mediated Communication* 12(4), 1409-1427.

- Sebba, Mark. 2013. "Multilingualism in written discourse". *International Journal of Bilingualism* 17(1), 97-118.
- Seidlhofer, Barbara. 2004. "Research perspectives on teaching English as a lingua franca". *Annual Review of Applied Linguistics* 24, 209-239.
- Seidlhofer, Barbara; Breiteneder, Angelika; Pitzl, Marie-Luise. 2006. "English as a lingua franca in Europe: Challenges for applied linguistics". *Annual Review of Applied Linguistics* 26, 3-34.
- Seidlhofer, Barbara. 2007. "English as a lingua franca and communities of practice". In Volk-Birke, Sabine; Lippert, Julia (eds.). *Anglistentag 2006 Halle Proceedings*. Trier: Wissenschaftlicher Verlag Trier, 307-318.
- Seidlhofer, Barbara. 2009a. "Accommodation and the idiom principle in English as a lingua franca". *Intercultural Pragmatics* 6(2), 195-215.
- Seidlhofer, Barbara. 2009b. "Orientations in ELF Research: Form and Function". In Mauranen, Anna; Ranta, Elina (eds.). *English as a Lingua Franca: Studies and findings*. Newcastle upon Tyne: Cambridge Scholars Publishing, 37-59.
- Seidlhofer, Barbara. 2011. *Understanding English as a Lingua Franca*. Oxford: Oxford University Press.
- Seidlhofer, Barbara; Widdowson, Henry G. 2007. "Idiomatic variation and change in English. The idiom principle and its realizations". In Smit, Ute; Dollinger, Stefan; Hüttner, Julia; Kaltenböck, Gunther; Lutzky, Ursula (eds.). *Tracing English through time. Explorations in language variation*. (Festschrift for Herbert Schendl, Austrian Studies in English vol. 95) Wien: Braumüller, 359-374.
- Sun, Hong-mei. 2010. "A study of the features of Internet English from the linguistic perspective". *Studies in Literature and Language*, 1(7), 98-103.
- Squires, Lauren. 2010. "Enregistering internet language". *Language in Society* 39(4), 457-492.
- Sveningsson Elm, Malin. 2009. "How Do Various Notions of Privacy Influence Decisions in Qualitative Internet Research?" In Baym, Nancy K.; Markham, Annette N. (eds.). *Internet Inquiry*. Thousand Oaks: SAGE Publications, 69-87.
- Thurlow, Crispin. 2006. "From statistical panic to moral panic: The metadiscursive construction and popular exaggeration of new media language in the print media". *Journal of Computer-Mediated Communication* 11(3), 667-701.
- Trudgill, Peter. 2002. *Sociolinguistic variation and change*. Edinburgh: Edinburgh University Press.
- Vettorel, Paola. 2014. *English as a Lingua Franca in Wider Networking: Blogging Practices*. Berlin: Walter de Gruyter.

- Viégas, Fernanda B. 2005 “Bloggers’ expectations of privacy and accountability: An initial survey”. *Journal of Computer-Mediated Communication* 10(3), <http://www.uvm.edu/pdodds/files/papers/others/2005/viegas2005a.pdf> (last accessed March 9, 2017).
- Walther, J. 1996. “Computer-mediated Communication: Personal, Interpersonal and Hyperpersonal Interaction”. *Communication Research* 23(1), 3-43.
- Walther, J. 2007. “Selective self-presentation in computer-mediated communication: Hyperpersonal dimensions of technology, language, and cognition”. *Computers in Human Behavior* 23(5), 2538-2557.
- Warschauer, Mark; El Said, Ghada R.; Zohry, Ayman. 2007. “Language Choice Online. Globalization and Identity in Egypt”. In Danet, Brenda; Herring, Susan C. (eds.). *The Multilingual Internet: Language, Culture, and Communication Online*. Oxford: Oxford University Press, 303-318.
- Widdowson, Henry G. 2007. *Discourse Analysis*. Oxford: Oxford University Press.
- Widdowson, Henry G. 2012. “ELF and the inconvenience of established concepts”. *Journal of English as a Lingua Franca* 1. 5–26.
- Widdowson, Henry G. 2015. “ELF and the pragmatics of language variation” *Journal of English as a Lingua Franca* 4 (2). 359-372.
- Wenger, Etienne; McDermott, Richard; Snyder, William M. 2002. *Cultivating communities of practice*. Cambridge M.A.: Harvard Business School Press.
- Wright, Sue. 2004. “Thematic Introduction”. *International Journal on Multicultural Societies* 6(1). 5-13.
- Yates, JoAnne; Orlikowski, Wanda.J. 1993. “Knee-jerk Anti-LOOPism and other E-mail Phenomena: Oral, Written, and Electronic Patterns in Computer-Mediated Communication”. MIT Sloan School Working Paper No. 3578-93 Presented at the 53rd Annual Meeting of the Academy of Management, Atlanta, GA, August 1993. <http://ccs.mit.edu/papers/CCSWP150.html>, last accessed April 21st 2017.

Internet sources:

- Ess, Charles and the AoIR ethics working committee. 2002. “Ethical decision-making and Internet research: Recommendations from the AoIR ethics working committee”. <http://www.aoir.org/reports/ethics.pdf> (last accessed March 20th 2017).
- <http://www.internetworldstats.com/stats7.htm>, last accessed April 1st 2017
- <http://knowyourmeme.com/memes/beautiful-cinnamon-roll-too-good-for-this-world-too-pure>, last accessed March 29th 2017.
- <https://techterms.com/definition/microblogging>, last accessed March 16th 2017.

https://www.tumblr.com/login?redirect_to=%2Fdashboard, last accessed March 16th 2017.

<https://unwrapping.tumblr.com/post/102742630432/tumblr-tags-limit>, last accessed April 3rd 2017

<https://unwrapping.tumblr.com/post/121385928727/tumblr-individual-tag-character-limit>, last accessed April 3rd 2017

Screenshots depicted in Figure 1: Layout of a post (page 30), Figure 2: Layout of a post labeled (page 30) and Figure 3: Reblogged post (page 31) taken from a password protected account created for the sake of this thesis on www.tumblr.com, last accessed April 3rd 2017).

Appendix

List of figures

Figure 1: Layout of a post	30
Figure 2: Layout of a post labeled	30
Figure 3: Reblogged post	31
Figure 4: Example 1	46
Figure 5: Example 2	48
Figure 6: Example 3	49
Figure 7: Example 4	50
Figure 8: Example 5	52
Figure 9: Example 6	53
Figure 10: Example 7	55
Figure 11: Example 8	56
Figure 12: Example 9	58
Figure 13: Example 10	58
Figure 14: Example 11	68
Figure 15: Example 12	70
Figure 16: Example 13	70

Abstract

The present thesis is concerned with a linguistic phenomenon on the website Tumblr. To be precise, the research conducted within this thesis investigates the communicative nature of tags as they are used on Tumblr. Furthermore, since much of the communication on the website is achieved by making use of ELF (English as a lingua franca), a second research interest explored in this thesis is the form this ELF interaction takes and the ELF concepts that come into play in this context.

In line with Androutsopoulos' discourse-centered online ethnography (DCOE, 2008), the analysis triangulates observation, log data, and interview data, and incorporates participants' own input regarding communicative tag usage and ELF on Tumblr. As a result, findings and conclusions drawn from the data are considered to be in line with the participants' experience of this phenomenon. In short, the present thesis claims that tags on Tumblr are not only used for organizational purposes or structuring of content, but are also used by bloggers to fulfill communicative purposes such as expressing emotion, adding commentary, or relating a post to the blogger. These different uses of tags are presented and discussed in the analysis. Judging from the log and interview data for this thesis, ELF is the most prominent language choice for tags for reasons of reach, community considerations, and context of the content shared. Several ELF-related concepts (among them formulaic expressions, communities of practice, and signaling of culture and/or identity) are found within the context of communicative tags and thus addressed in the analysis.

It is thereby my hope that the present thesis contributes to the fields of Internet linguistics and ELF by providing insights into a linguistic phenomenon realized through ELF in written computer-mediated communication.

Abstract (German)

Die vorliegende Diplomarbeit beschäftigt sich mit einem linguistischen Phänomen auf der Website Tumblr. Genauer gesagt untersucht die hier durchgeführte Forschung die kommunikative Natur von Tags wie sie auf Tumblr verwendet werden. Da ein großer Anteil der Kommunikation auf dieser Website durch ELF (‘English as a lingua franca’, Englisch als Kontaktsprache) erreicht wird, ist das zweite Forschungsinteresse in dieser Diplomarbeit die Form die diese ELF-Interaktion annimmt und die Verbindung zu ELF-Konzepten welche in diesem Zusammenhang relevant sind.

In Einklang mit Androutsopoulos’ discourse-centered online ethnography (‘Diskurszentrierte Online-Ethnographie’, 2008) verbindet die Analyse Beobachtungen, Protokolldaten, und Interviewdaten und inkorporiert zusätzlich Beiträge zur Verwendung von Tags und ELF von Beteiligten Tumblr-Benutzern selbst. Aus diesem Grund wird angenommen, dass die aufgrund der Analyse geschlossenen Schlussfolgerungen und Ergebnisse den Erfahrungen der Beteiligten entsprechen. Zusammengefasst stellt die vorliegende Diplomarbeit die Behauptung auf, dass Tags auf Tumblr nicht nur für organisatorische Zwecke oder zur Strukturierung von Inhalten verwendet werden, sondern darüber hinaus auch von Tumblr-Benutzern eingesetzt werden um kommunikative Zwecke (wie das Ausdrücken von Emotionen, Hinzufügen von Kommentaren, oder Beziehen eines Blogeintrags auf sich selbst) zu erfüllen. Diese unterschiedlichen Verwendungsweisen von Tags werden in der Analyse dargestellt und besprochen. Von den Protokoll- und Interviewdaten dieser Diplomarbeit schließend stellt sich ELF als hervorstechende Sprachwahl für Tags heraus und wird aufgrund von Faktoren wie Reichweite, Gemeinschaftsgefühlen und Kontext der geteilten Inhalte anderen Sprachen gegenüber bevorzugt. Im Kontext von kommunikativen Tags finden sich eine Reihe von mit ELF im Zusammenhang stehenden Konzepten (darunter formelhafte Ausdrücke, sogenannte ‚communities of practice’ oder ‚Praxisgemeinden’, und das Signalisieren von Kultur und/oder Identität), welche aus diesem Grund in der Analyse besprochen werden.

Insgesamt ist es meine Hoffnung, dass die im Rahmen dieser Diplomarbeit durchgeführte Forschung nützliche Erkenntnisse zu den Forschungsgebieten der

Internetlinguistik und Englisch als lingua franca beiträgt indem sie das Phänomen der durch ELF realisierten kommunikativen Tags in schriftlicher computervermittelter Kommunikation auf Tumblr untersucht.